

541 TGGATAGCTTTGACACGACCCCTTTATGTACAAAAATGCGAGGTTCTTACCTCTACT 600
541 TGGATAGCTTTGACACGACCCCTTTATGTACAAAAATGCGAGGTTCTTACCTCTACT 600
601 TACACACGAACTGGGCTGACGAGAGATTTGGAAGCATGTAACATTGGCTCGGTAC 660
601 TACAAACGAACTGGGCTGACGAGAGATTTGGAAGCATGTAACATTGGCTCGGTAC 660
661 TCAGATCTTCAGAGAGAGAGGCTTGGAAAACTCTCAATCTTGAAGAAAGGCTCCAA 720
661 TCAGATCTTCAGAGAGAGAGGCTTGGAAAACTCTCAATCTTGAAGAAAGGCTCCAA 720
721 CCTACTAATAGATCATATTTCTCGGTTGTTCAACATCTTACACAGAAATGATCACTG 780
721 CCTACTAATAGATCATATTTCTCGGTTGTTCAACATCTTACACAGAAATGATCACTG 780
781 TTACGTACTGGCATCTTCCAAACGTGTTCCACTTGAAGAAAGGCTTAACTTCAACAGT 840
781 TTACGTACTGGCATCTTCCAAACGTGTTCCACTTGAAGAAAGGCTTAACTTCAACAGT 840
841 AGATGGGACCATTTGTGAGTGTGAGGATGTCATCAAAAAGATTAACATCAAGCCCA 900
841 AGATGGGACCATTTGTGAGTGTGAGGATGTCATCAAAAAGATTAACATCAAGCCCA 900
841 AGATGGGACCATTTGTGAGTGTGAGGATGTCATCAAAAAGATTAACATCAAGCCCA 900
901 GGAATTAACGATTAAGTGAAGATTTGGGCTCACAATGATCGGAGGTTCTTGAAT 960
901 GGAATTAACGATTAAGTGAAGATTTGGGCTCACAATGATCGGAGGTTCTTGAAT 960
961 TGCATAATAGGTTGAAGATTTGGGCTCACAATGATCGGAGGTTCTTGAAT 960
961 TGCATAATAGGTTGAAGATTTGGGCTCACAATGATCGGAGGTTCTTGAAT 960
1021 CCAGCCACTTTGGGATGAGTGAAGGATTTCTGGCACTGAGCTTATGTTGATGAC 1080
1021 CCAGCCACTTTGGGATGAGTGAAGGATTTCTGGCACTGAGCTTATGTTGATGAC 1080
1081 GCACAAAACTATTTGGGCTCAACCAAGATTTGTCATGATGATGAGCAGAAAG 1140
1081 GCACAAAACTATTTGGGCTCAACCAAGATTTGTCATGATGATGAGCAGAAAG 1140
1141 AATACTAACAAATGAGATCTATCTATTAACAGTGTGAGGCTTTTCCAGGTGG 1200
1141 AATACTAACAAATGAGATCTATCTATTAACAGTGTGAGGCTTTTCCAGGTGG 1200
1201 GCGGCTGAACATCTGCGGCTTGAACGAGAAAGATGAGGCTGGGAGGCACT 1260
1201 GCGGCTGAACATCTGCGGCTTGAACGAGAAAGATGAGGCTGGGAGGCACT 1260
1261 CTTACTATGAGGCTGCTGGGCTTCAAGAACCCAGAAATCACTCCATCTACAAAGAG 1320
1261 CTTACTATGAGGCTGCTGGGCTTCAAGAACCCAGAAATCACTCCATCTACAAAGAG 1320
1321 CTTGATGAGGCTGCTGGGCTTCAAGAACCCAGAAATCACTCCATCTACAAAGAG 1380
1321 CTTGATGAGGCTGCTGGGCTTCAAGAACCCAGAAATCACTCCATCTACAAAGAG 1380
1381 CTTGATGAGGCTGCTGGGCTTCAAGAACCCAGAAATCACTCCATCTACAAAGAG 1440
1381 CTTGATGAGGCTGCTGGGCTTCAAGAACCCAGAAATCACTCCATCTACAAAGAG 1440
1441 ACTGTCAAAACCGGCTATTTCAATGAGCGGATGAGAGCATCTGCTGGCTTACAG 1500
1441 ACTGTCAAAACCGGCTATTTCAATGAGCGGATGAGAGCATCTGCTGGCTTACAG 1500
1501 CAAGAACTGAAAGATGCTGAGCGGAAAGATCAAGAAAGCCCTGCTC 1560
1501 CAAGAACTGAAAGATGCTGAGCGGAAAGATCAAGAAAGCCCTGCTC 1560
1561 CCTGAATAGAAAGAGCGGTAGAGGAGAGATGAGATCTGATTAATGAGAAAGGCAAGA 1620
1561 CCTGAATAGAAAGAGCGGTAGAGGAGAGATGAGATCTGATTAATGAGAAAGGCAAGA 1620
1621 GCAAGTACGCTGAGAGACACAGAGAGACATCAAGGTGACAAAGTTACCCAGGCGAAGAG 1680

1621 GCAGGTACGCTGAGAGACACACAGAGAGACATCAAGGTTGACAAAGTTACCCAGGCAAGAG 1680
1681 AAGATTTGGCTTTAGCTATCTTTCAACCCAGGCGGATTTGAATAGGAAATCTGGCG 1740
1681 AAGATTTGGCTTTAGCTATCTTTCAACCCAGGCGGATTTGAATAGGAAATCTGGCG 1740
1741 TGTATCCACCAATGGCGGAAACAAAGTATGATCTCAAAAGTATGAGGCAAGAGAG 1800
1741 TGTATCCACCAATGGCGGAAACAAAGTATGATCTCAAAAGTATGAGGCAAGAGAG 1800
1801 TACAAAGTGAAGCATACCAAGGTAAGTCAATTTGAACAGAAAGGAGCGGCTCTGTT 1860
1801 TACAAAGTGAAGCATACCAAGGTAAGTCAATTTGAACAGAAAGGAGCGGCTCTGTT 1860
1861 CAAGCTTCAGGCTTGAAGTGAAGGCGCTACGCTTTTCAAGAGAGGAGTTGCTGA 1920
1861 CAAGCTTCAGGCTTGAAGTGAAGGCGCTACGCTTTTCAAGAGAGGAGTTGCTGA 1920
1921 AACAGATACCTGACCAATCAATCAACGAGAGGCTTAAACATGACGAAAGTAC 1980
1921 AACAGATACCTGACCAATCAATCAACGAGAGGCTTAAACATGACGAAAGTAC 1980
1981 TATAAGATCTGTAAGATCTCAGAGACACAGATCAGAAATGCTTGTGATATTGAACG 2040
1981 TATAAGATCTGTAAGATCTCAGAGACACAGATCAGAAATGCTTGTGATATTGAACG 2040
2041 AAGTGTGTAAGCGGAAAGAGAGGAGTCCCTGTCCTTCAACGAGTATGATATCA 2100
2041 AAGTGTGTAAGCGGAAAGAGAGGAGTCCCTGTCCTTCAACGAGTATGATATCA 2100
2101 CCATTTCAAGATTTGCGTACGAGATCTCAAGACACGAGCAGAGCCTTCAAAAGTC 2160
2101 CCATTTCAAGATTTGCGTACGAGATCTCAAGACACGAGCAGAGCCTTCAAAAGTC 2160
2161 CCAACCATCGAGTCTATGAGTGTGAGGATTCAGGTTCAAGTATCTGAATCATCAAAAG 2220
2161 CCAACCATCGAGTCTATGAGTGTGAGGATTCAGGTTCAAGTATCTGAATCATCAAAAG 2220
2221 GTGACTAAGAAAGATCTGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2280
2221 GTGACTAAGAAAGATCTGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2280
2281 GATGTAAAGAGATGAGAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2340
2281 GATGTAAAGAGATGAGAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2340
2341 AATGGGTTAAGCAACCCGTTTAACTCTGTATCATTTGATGAGGCAATTTGCCCTGCAATCA 2400
2341 AATGGGTTAAGCAACCCGTTTAACTCTGTATCATTTGATGAGGCAATTTGCCCTGCAATCA 2400
2401 GGGAGCTGCTGGGACCTGATTTGCAATCTGCAAACTTAAAGAAAGTGTGTTGGCGGAGC 2460
2401 GGGAGCTGCTGGGACCTGATTTGCAATCTGCAAACTTAAAGAAAGTGTGTTGGCGGAGC 2460
2461 CCAAAACAAATGCGGCTTTTAACTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2520
2461 CCAAAACAAATGCGGCTTTTAACTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2520
2521 TGCATGAGTGTACATTAAGATCTTGAAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2580
2521 TGCATGAGTGTACATTAAGATCTTGAAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2580
2581 GTCTTCAGCTCTTCTTACAGCAAGAGGATGAAAGCGGTAAACCAATGTGCTGATTAATTC 2640
2581 GTCTTCAGCTCTTCTTACAGCAAGAGGATGAAAGCGGTAAACCAATGTGCTGATTAATTC 2640
2641 ATCATAGTATACCAAGGAGCAACAAAGCGGACCAAAAGTGTGATTTCTAAGCTGTTTC 2700
2641 ATCATAGTATACCAAGGAGCAACAAAGCGGACCAAAAGTGTGATTTCTAAGCTGTTTC 2700
2701 AGAGATGGGTGAACAGCTACAGATTGATCTCAAAATTCAGAAATCATGACTGCGGCT 2760

Db 2701 AGAGGATGGGTGAAGACAGCTACAGATTGACTACAAATAATCATGATGCGGCT 2760
Qy 2761 GCATCGCAAGGACTTACGCGGAAAGCGGTATATGCTGACAGTACAAAGTCAAGAGAT 2820
Db 2761 GCATCGCAAGGACTTACGCGGAAAGCGGTATATGCTGACAGTACAAAGTCAAGAGAT 2820
Qy 2821 CCACCTACTGCGAGACTTGTGACACGCTGAACGCTTACTTACACGACGAGAAAACGC 2880
Db 2821 CCACCTACTGCGAGACTTGTGACACGCTGAACGCTTACTTACACGACGAGAAAACGC 2880
Qy 2881 ATTGTCTGGAAGACGCTAGTGTGATCCCTGATTAAGACACTTACAGCTAAATATCC 2940
Db 2881 ATTGTCTGGAAGACGCTAGTGTGATCCCTGATTAAGACACTTACAGCTAAATATCC 2940
Qy 2941 GGGGATTTACGGCTTCAATTGAGACCTGACGCGGCAACGAGCCATTAATGCGACGC 3000
Db 2941 GGGGATTTACGGCTTCAATTGAGACCTGACGCGGCAACGAGCCATTAATGCGACGC 3000
Qy 3001 GTTCTGTATAGCCGACAGACAGCTGATGTCTCCAGATTAAGTGAACGCTGCGTGGGCG 3060
Db 3001 GTTCTGTATAGCCGACAGACAGCTGATGTCTCCAGATTAAGTGAACGCTGCGTGGGCG 3060
Qy 3061 AAGGCTTTAGACCAAGCTCTTGCCCAACGCTGCTGACGAGACAGCAGTGGAG 3120
Db 3061 AAGGCTTTAGACCAAGCTCTTGCCCAACGCTGCTGACGAGACAGCAGTGGAG 3120
Qy 3121 ACCTTGCACCCATTCACAGATGACAGACGCTACTCACTGAAATGCGACTGACTTT 3180
Db 3121 ACCTTGCACCCATTCACAGATGACAGACGCTACTCACTGAAATGCGACTGACTTT 3180
Qy 3181 TGCAACAGGTTCTTTGAGTAGACTGACAGTGGTTATTTCCGCTCTACCGCTGCA 3240
Db 3181 TGCAACAGGTTCTTTGAGTAGACTGACAGTGGTTATTTCCGCTCTACCGCTGCA 3240
Qy 3241 CTTAATTACAGGATCAGCACTGCGATTAACCTGCCAGGAAAGAACTGATGGCTTAT 3300
Db 3241 CTTAATTACAGGATCAGCACTGCGATTAACCTGCCAGGAAAGAACTGATGGCTTAT 3300
Qy 3301 AGAAGGTAAGCAAGGATTTGTCAAGGAGATTCCTGTCATCAAAAGCGGTTGACCA 3360
Db 3301 AGAAGGTAAGCAAGGATTTGTCAAGGAGATTCCTGTCATCAAAAGCGGTTGACCA 3360
Qy 3361 GGGAGGGTATGATATATAGGAATATATCAATCAAGACTCTCCCAAAATTAATGTG 3420
Db 3361 GGGAGGGTATGATATATAGGAATATATCAATCAAGACTCTCCCAAAATTAATGTG 3420
Qy 3421 GTTCCATTAAATCGCCGGTTCCTGATCTGTTGATCTTGAACCAAAAGGAGGATCA 3480
Db 3421 GTTCCATTAAATCGCCGGTTCCTGATCTGTTGATCTTGAACCAAAAGGAGGATCA 3480
Qy 3481 ACTGATCAGACGGATTCCTATCTAATGTAAGGCAAACTGTGTGTGATCGGCGAT 3540
Db 3481 ACTGATCAGACGGATTCCTATCTAATGTAAGGCAAACTGTGTGTGATCGGCGAT 3540
Qy 3541 CCTATCAGCATTTCCAGGAAAGAAAGTAGATGATGATGATGATGATGATGATGATG 3600
Db 3541 CCTATCAGCATTTCCAGGAAAGAAAGTAGATGATGATGATGATGATGATGATGATG 3600
Qy 3601 AGGTGTGATCTGATTTGGGAATACCTAGCCATGTGCGTAATATGATCATTTATCT 3660
Db 3601 AGGTGTGATCTGATTTGGGAATACCTAGCCATGTGCGTAATATGATCATTTATCT 3660
Qy 3661 AATGTTAGACCCCGTACAGGAAACCATCACTACCAAGTGCAGGATCAGCTATTCAC 3720
Db 3661 AATGTTAGACCCCGTACAGGAAACCATCACTACCAAGTGCAGGATCAGCTATTCAC 3720
Qy 3721 CACAGCATGCTAACGTTAAGGCTGTCCACCACTGGAACACTGGCGAAACATGTGTGCT 3780
Db 3721 CACAGCATGCTAACGTTAAGGCTGTCCACCACTGGAACACTGGCGAAACATGTGTGCT 3780
Qy 3781 ATAAGGTAATGGGCTTGTGATCGCGCAACCGAATATCATCATCTGCGGTGGACGCTCA 3840
Db 3781 ATAAGGTAATGGGCTTGTGATCGCGCAACCGAATATCATCATCTGCGGTGGACGCTCA 3840

Qy 3841 TTTAGGTTTACCCGCTGTCTGTACAGCTTACAGAACTGCGGAAATATAGAGTTCTCTTC 3900
Db 3841 TTTAGGTTTACCCGCTGTCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTTC 3900
Qy 3901 GTGTCTTTCGGAAGGACCAACCGCAACCAACATGACAGGACAGACTCGGTGTAGTG 3960
Db 3901 GTGTCTTTCGGAAGGACCAACCGCAACCAACATGACAGGACAGACTCGGTGTAGTG 3960
Qy 3961 CTTGACAACTCTATCAAGGCTCAACAGTACAGGACAGGAGAGTCTCAGGCTACAGA 4020
Db 3961 CTTGACAACTCTATCAAGGCTCAACAGTACAGGACAGGAGAGTCTCAGGCTACAGA 4020
Qy 4021 GTGATCAGAGTGCATTTACCAAGAGGCTGACCAAGCTATGCTTAATGCTTAATAGC 4080
Db 4021 GTGATCAGAGTGCATTTACCAAGAGGCTGACCAAGCTATGCTTAATGCTTAATAGC 4080
Qy 4081 AAAGGTCACAGGTTCCGGAAGTGTGCGGTGCACTGTATCCGAAATGCGCGCTGCTTT 4140
Db 4081 AAAGGTCACAGGTTCCGGAAGTGTGCGGTGCACTGTATCCGAAATGCGCGCTGCTTT 4140
Qy 4141 GATGACAGCCAAATAGCTGTGCGGACCGCTAGACTTGTGAAGCAACGCTCATATA 4200
Db 4141 GATGACAGCCAAATAGCTGTGCGGACCGCTAGACTTGTGAAGCAACGCTCATATA 4200
Qy 4201 CATGCTGTAGAGCCCAATTTTCTAAGATGCGGGAACCGAAGGCGCACTTAAGCTGCA 4260
Db 4201 CATGCTGTAGAGCCCAATTTTCTAAGATGCGGGAACCGAAGGCGCACTTAAGCTGCA 4260
Qy 4261 GCTGCTTACATGACATAGCGCTCATGTCACAGCTGAGCGGATTAACAAATATACGTA 4320
Db 4261 GCTGCTTACATGACATAGCGCTCATGTCACAGCTGAGCGGATTAACAAATATACGTA 4320
Qy 4321 CCGCTACTGTCAACCGGCACTTATCTGTGTGCAAAATGAGTGAATGCAATCTTCAT 4380
Db 4321 CCGCTACTGTCAACCGGCACTTATCTGTGTGCAAAATGAGTGAATGCAATCTTCAT 4380
Qy 4381 CACCTGTTCACGCTTTCGACACTAGCGGATGCGGATGTCACATATATCTGATATAA 4440
Db 4381 CACCTGTTCACGCTTTCGACACTAGCGGATGCGGATGTCACATATATCTGATATAA 4440
Qy 4441 CAATGGAGACCAAGATTAATGAGGCTTACCGGCAAAAGCGTGAATTTCTGAT 4500
Db 4441 CAATGGAGACCAAGATTAATGAGGCTTACCGGCAAAAGCGTGAATTTCTGAT 4500
Qy 4501 GATGACAGCCAGTACATTTGATCTGTGAGGCTTCAACCAACAGCTCTTTGGCAGGC 4560
Db 4501 GATGACAGCCAGTACATTTGATCTGTGAGGCTTCAACCAACAGCTCTTTGGCAGGC 4560
Qy 4561 AGACCAAGTTACTCCGTCATGAGGCGCAATGATTAATCATACGTAAGGATGACGATTC 4620
Db 4561 AGACCAAGTTACTCCGTCATGAGGCGCAATGATTAATCATACGTAAGGATGACGATTC 4620
Qy 4621 CATCAGACCGCAAGACATTTGCGGAATCATATGATGATGATGATGATGATGATGATG 4680
Db 4621 CATCAGACCGCAAGACATTTGCGGAATCATATGATGATGATGATGATGATGATGATG 4680
Qy 4681 AATGAGCAATTTGCTGTGATCTCTGAGGAGATGATGATGATGATGATGATGATGATG 4740
Db 4681 AATGAGCAATTTGCTGTGATCTCTGAGGAGATGATGATGATGATGATGATGATGATG 4740
Qy 4741 CCAAGTAGAGGATGAGGAGGCTGCTCAACCTGACACACTGCTGCTGCTGCTGCTGCT 4800
Db 4741 CCAAGTAGAGGATGAGGAGGCTGCTCAACCTGACACACTGCTGCTGCTGCTGCTGCT 4800
Qy 4801 GCTATGACGCGTGAAGCGCTATACAGGTTGCGCTCTGCAAGAAAGAAAGTGTGCGGTA 4860
Db 4801 GCTATGACGCGTGAAGCGCTATACAGGTTGCGCTCTGCAAGAAAGAAAGTGTGCGGTA 4860
Qy 4861 TGCTCATCATTTCTGTTGCGGAAGTACAGATTCACAGGCTGCGAAAGCTTACATGACGC 4920
Db 4861 TGCTCATCATTTCTGTTGCGGAAGTACAGATTCACAGGCTGCGAAAGCTTACATGACGC 4920

Qy 4921 AACAGCTCTGTTTTCAGGCGCTGCTACACCGCGCTGTACACCCAGGAGTAGCGGAA 4980
|||||
Db 4921 AACAGCTCTGTTTTCAGGCGCTGCTACACCGCGCTGTACACCCAGGAGTAGCGGAA 4980
|||||
Qy 4981 ATAAATCTAGAAACGCCACC 5000
|||||
Db 4981 ATAAATCTAGAAACGCCACC 5000
|||||
RESULT 2
LOCUS MEU01065 4471 bp RNA linear VRL 01-JUN-1995
DEFINITION Western equine encephalomyelitis virus nonstructural polyprotein
gene, partial cds.
ACCESSION U01065.1 GI:393033
VERSION nsp1; nsp2; nsp3; nsp4.
KEYWORDS Western equine encephalomyelitis virus
SOURCE Western equine encephalomyelitis virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; MEV complex.
1 (bases 1 to 4471)
REFERENCE Weaver, S.C., Hagenbaugh, A., Bellet, L., Netesov, S.V., Volchkov, V.E.,
AUTHORS Chang, G.-J., Clarke, D.K., Goussard, L., Scott, T.W., Trent, D.W. and
Holland, J.J.
TITLE A comparison of the nucleotide sequences of eastern and western
equine encephalomyelitis viruses with those of other alphaviruses
and related RNA viruses
JOURNAL Virology 197 (1), 375-390 (1993)
MEDLINE 94025587
PubMed 8105605
REFERENCE 2 (bases 1 to 4471)
AUTHORS Weaver, S.C.
TITLE Direct Submision
JOURNAL Submitted (25-AUG-1993) Scott C. Weaver, Biology, University of
California, San Diego, La Jolla, CA 92037-0116 USA
FEATURES
source
1. 4471
/organism="Western equine encephalomyelitis virus"
/mol_type="genomic RNA"
/isolate="5614"
/db_xref="taxon:11039"
/clone="83, 34, 44, 75, 312"
/note="Western equine encephalitis virus"
<1..>4471
/note="Termination codon at position 4253..4255"
/codon_start=2
/transl_except="(pos:4253..4255,aa:OTHER)
/product="nonstructural polyprotein"
/protein_id="AA67651.1"
/db_xref="GI:393034"
/translation="TIRKVPVAFDSFVPLRTSHGPDIGLFRSRLLLEPTVKAPAI
TMADVEHRLGQAEVEVAABEIRALPPLPEIKETVAEDLINOAGASVET
PRGIRVTSYQGEKEIGSVALISPOAVNLSKICINHLAEVOVMTKGAGHVK
PYHKVIVPEGTAVPOFQALSSATVFNREFFVNLHIIANGALTDDEYK
TVKTQDDEYVFDIADKCYKREDAEPLCTGLVDPPHPPAVESLKTNPAA
PLTVGVGPGSGKGIISAVTKKDIIVASAKENCAETIRPVRBRDVAAPVDS
LNLKAPVNTLVIDEAFACAGTILALIALVKKPVLCDDPKCGCFNNMCKVH
NHDICTEVYHSISRCTQVTAIVSTLFYDKRMKTVNCPADKIIDTGTGTHKID
LILTCFRGVNQQLIDYKNHEIMTAASQGLTRGVAVRKYVENPLYSOTSHVY
LVRTKRIIVKTLADGPWIKTPTAKYGGDTASLDIDRSHDAIMARVLDKPTADY
FONKLVNCAKALBPVLATANIIVLTROOMETLHPKHBRAYSPEALNFCPTGAVD
LDGLFSAPATVATLYRDQHMNSPKNWYGNREVAKEISRPYCITKATVDTGAVDI
RMTIKOYSPITNVVPLNRPLPSLIDYHKQGTTHDSGFSKANGSVLITGPISI
PKRVESWGPLPTVIRCDLDLGIPIHVGRKDIILFVNRTYKVNHYQCCDHAIIHS
MLTCAVHNLVTGTCVAILGYGLADRAENITAVRSFRTRCOPNOMTEVLF
VFPRKDNHNDHODRLGVLDNIYQGSTREAGRAPYRVIRGDISKADQALVNA
NSKQPGSGCGALYKRPAPAFROPIAVGTARLVKSEPLIIHAVGPMFSQMBE
LKLAAYMSIASIVNAERTIKIVSLSTGIVSGKRVMSIHLHPTAPPTADYT
IYCDKQWERTITRAIRKESVEIIDDKVDVLDVAVHNSLSLAGRPGVSDGKT
YSIEGRFHOTADIAETIHAMVPKKSANEOICLYLIGESMSIRSKCPVEESBA
PPHTPCICNVAETAEVRLRSKKEFAFACSSFPKXYITGVOKQCKPVLFGS
VPPAVHPRKYAEIILETPSPPTTTTVCPETVBERLPSPIBSAPSESLISFGVS

Query Match 70.7%; Score 3536; DB 14; Length 4471;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 3594; Conservative 0; Mismatches 75; Indels 3; Gaps 1;
mat_peptide
/product="nsp1 protein"
mat_peptide
/product="nsp2 protein"
mat_peptide
/product="nsp3 protein"
mat_peptide
/product="nsp4 protein"
BASE COUNT 1272 a 1110 c 1116 g 973 t
ORIGIN
FSSATNSSTAMSDYDRFVVTADVHOANTSWSIPSPAGLDVOLPSPDTSHWSIPS
ALGEVTPSVODLTAECAERGLAEIMODENTAPFOLSHRPVAPARRPIPSRS
TASAPVPRPTRYOOPGVARAISELEBEYIRHONKXRYBGAIFYSETGCHL
QOKSROCKLODPIFERAVHEKTYAPRLDLREHLOKQLOLCASEGRSRYSQS"
Qy 1332 AACATTAGAAAGTACCTGCGCTTGTGACTATTGTGATTCGACGCTTACGAGCA 1391
|||||
Db 1 AACATTAGAAAGTACCTGCGCTTGTGACTATTGTGATTCGACGCTTACGAGCA 60
1392 CGGCTCGATATGGGCTTCCGCCGTGAGCTCAAGCTGCTGTGAACCACTGTCAAAC 1451
|||||
Db CGGCTCGATATGAGGCTTCCGCCGTGAGCTCAAGCTGCTGTGAACCACTGTCAAAC 120
1452 CGCACCGCTATTACATATGCCGATGTGAGACATCTGTGGCTTACAGCAAGAGCTGA 1511
|||||
Db CGCACCGCTATTACATATGCCGATGTGAGACATCTGTGGCTTACAGCAAGAGCTGA 180
1512 AGAAGTGTGCTGAGGGAAGATCAGAGAAAGCCCTGCACCCCTTGCTCCCTGAAATGA 1571
|||||
Db AGAAGTGTGCTGAGGGAAGATCAGAGAAAGCCCTGCACCCCTTGCTCCCTGAAATGA 240
1572 AAAAGAGACCGTAGAGCAGAAAGTGAACCTATTATGCAAGAGCAGAGCAGTAGCGT 1631
|||||
Db AAAAGAGACCGTAGAGCAGAAAGTGAACCTATTATGCAAGAGCAGAGCAGTAGCGT 300
1632 GGAAGACACGAGAGACATCAGGCTGACAAATGCCAGCCGAAAGAAATTGGGTC 1691
|||||
Db GGAAGACACGAGAGACATCAGGCTGACAAATGCCAGCCGTAAGAAAGAAATTGGGTC 360
1692 TTAGCTATATTTCACCCGAGGCGGTATTAATGAGAAACCTGGGTATTCACACC 1751
|||||
Db TTAGCTATATTTCACCCGAGGCGGTATTAATGAGAAACCTGGGTATTCACACC 420
1752 ATGGCGGAACAAGTACTGTATATGACTCACAAAGGTAGGCGAGAGATCAAAAGTGA 1811
|||||
Db ATGGCGGAACAAGTACTGTATATGACTCACAAAGGTAGGCGAGAGATCAAAAGTGA 480
1812 GCCATACACCGTAGAGCTATTGTACAGAAAGGACGCGGCTCCCTGTTCAAGACTTCA 1871
|||||
Db GCCATACACCGTAGAGCTATTGTACAGAAAGGACGCGGCTCCCTGTTCAAGACTTCA 540
1872 GGCATTGAGTGAAGCGCTACATCGTTTAAAGAGGAGAGTTCGTAACAGTACT 1931
|||||
Db GGCATTGAGTGAAGCGCTACATCGTTTAAAGAGGAGAGTTCGTAACAGTACT 600
1932 GCACCAATCCCAATCAACGAGAGGCGTAAACACTGACGAGAGTACTTAAGACTGT 1991
|||||
Db GCACCAATCCCAATCAACGAGAGGCGTAAACACTGACGAGAGTACTTAAGACTGT 660
1992 AAAAGCTCAGACACAGACTAGATAGTCTTGATATTGACGACGAAAGTGTGTA 2051
|||||
Db AAAAGCTCAGACACAGACTAGATAGTCTTGATATTGACGACGAAAGTGTGTA 720
2052 GCGAAGAGAGCAGGCTCTTGCTGCTTAAACCGGATCTGGTATATCAACATTCACGA 2111
|||||
Db GCGAAGAGAGCAGGCTCTTGCTGCTTAAACCGGATCTGGTATATCAACATTCACGA 780
2112 GTTGGCTGAGAGTCTCAAGACAGCAGCAGCACTTCAAAAGTCCCAACATCGG 2171
|||||

Db 781 GTTTCGACGAGAGTCTCAAGACACGACGACACCTCAAGAAAGTCCCAACCATCGG 840
Qy 2172 AGTTATAGAGTGCAGAGTTAGGTAATCTGGAAATCATCAAAAGCGTGTGACTAAGAA 2231
Db 841 AGTTATAGAGTGCAGAGTTAGGTAATCTGGAAATCATCAAAAGCGTGTGACTAAGAA 900
Qy 2232 AGATCTGGTGTGAGTGCGAAGAAAGAAACCTGGCGCAAAATCATCAAGGATGTAAAGAG 2291
Db 901 AGATCTGGTGTGAGTGCGAAGAAAGAAACCTGGCGCAAAATCATCAAGGATGTAAAGAG 960
Qy 2292 GATGAGACGTATGAGATGTCTGCTAGAGACTGCTGATTCAGTGGCTTCTAAATGGGGTTAA 2351
Db 961 GATGAGACGTATGAGATGTCTGCTAGAGACTGATTCAGTGGCTTCTAAATGGGGTTAA 1020
Qy 2352 GCACCCCGTTAACTCTGTATCATTTGATGAGGCAATTTGCCCTGCAATGACGAGGAGCTGCT 2411
Db 1021 GCACCCCGTTAACTCTGTATCATTTGATGAGGCAATTTGCCCTGCAATGACGAGGAGCTGCT 1080
Qy 2412 GGCACGTATGTCATTCGTCGCAAACTTAAAGAAAGTGTATTTGGCGGGGACCCCAAAACAATG 2471
Db 1081 GGCACGTATGTCATTCGTCGCAAACTTAAAGAAAGTGTATTTGGCGGGGACCCCAAAACAATG 1140
Qy 2472 CGGCTTCTTTAATCATGATGTGCTGAAAGTAACTTTTAAACCATGACATATGCACTGAAGT 2531
Db 1141 CGGCTTCTTTAATCATGATGTGCTGAAAGTAACTTTTAAACCATGACATATGCACTGAAGT 1200
Qy 2532 GTACATATTAAGCATCTCTAGAGAGTGTGACACAGACTGTAAACCGCCATGCTCTCAAGCT 2591
Db 1201 GTACATATTAAGCATCTCTAGAGAGTGTGACACAGACTGTAAACCGCCATGCTCTCAAGCT 1260
Qy 2592 CTTTCAAGACAAAGGAATGAAGACGGTTAACCAATGTCTGATTAATCATCATAGATAC 2651
Db 1261 CTTTCAAGACAAAGGAATGAAGACGGTTAACCAATGTCTGATTAATCATCATAGATAC 1320
Qy 2652 CACAGGACCAAAAGCCGCAAAAGATGTATCTGATTTCACTGTTTCAGAGAGTGGGT 2711
Db 1321 CACAGGACCAAAAGCCGCAAAAGATGTATCTGATTTCACTGTTTCAGAGAGTGGGT 1380
Qy 2712 GAACACGCTACAGATGTACTCAAAAAATCAGAAATCTGACTGGGTGTGATGCGAAGG 2771
Db 1381 GAACACGCTACAGATGTACTCAAAAAATCAGAAATCTGACTGGGTGTGATGCGAAGG 1440
Qy 2772 ACTTACGCGGAAGACGGTTATGCTGTAGAGTAAAGTCAACGAGAAATCCACTTACTC 2831
Db 1441 ACTTACGCGGAAGACGGTTATGCTGTAGAGTAAAGTCAACGAGAAATCCACTTACTC 1500
Qy 2832 GCACACTTCTGAGACGCTGAACGTGTATCTTACACGACAGAAAAACGATTTGTCTGAA 2891
Db 1501 GCACACTTCTGAGACGCTGAACGTGTATCTTACACGACAGAAAAACGATTTGTCTGAA 1560
Qy 2892 GACGCTAGCTGTGATCCCTGGATTAAGACCTTACAGCTAAATATCCCGGGGATTTTAC 2951
Db 1561 GACGCTAGCTGTGATCCCTGGATTAAGACCTTACAGCTAAATATCCCGGGGATTTTAC 1620
Qy 2952 GGCTTCAATGACGCTGGACGCGGAAACAGACGCCCTTATGGACGCTTCTTGAATTA 3011
Db 1621 GGCTTCAATGACGCTGGACGCGGAAACAGACGCCCTTATGGACGCTTCTTGAATTA 1680
Qy 3012 GCGCCAGACAGCTGATGTGTCCAGAAATAGGTGAACGTCTGCTGGCGGAAGGCTTTAGA 3071
Db 1681 ACCGACAGACAGCTGATGTGTCCAGAAATAGGTGAACGTCTGCTGGCGGAAGGCTTTAGA 1740
Qy 3072 GCGCAGTCTTGGCCACGCGCAACATTTGTCTGACGAGACAGAGTGGAGAGCTTGCACCC 3131
Db 1741 GCGCAGTCTTGGCCACGCGCAACATTTGTCTGACGAGACAGAGTGGAGAGCTTGCACCC 1800
Qy 3132 ATTGAAGATGACGAGAGCTACTCACTGAATGGCACTGAATCTTTTGCACAGGTT 3191
Db 1801 ATTGAAGATGACGAGAGCTACTCACTGAATGGCACTGAATCTTTTGCACAGGTT 1860
Qy 3192 CTTTGAAGTACCTGACAGTGGGTTATTTTCCGCTCTTACCGTGCACCTTACTTACG 3251
Db 1861 CTTTGAAGTACCTGACAGTGGGTTATTTTCCGCTCTTACCGTGCACCTTACTTACG 1920

Qy 3252 GGAATCAGACTGGGATTAATCGCCAGGGAGAAACATGTATGGGCTTAATAGAGGTAGC 3311
Db 1921 GGAATCAGACTGGGATTAATCGCCAGGGAGAAACATGTATGGGCTTAATAGAGGTAGC 1980
Qy 3312 AAAGAGTGTCAACGGGATATCCGTGCATCACAAAAGCGGTTGACACAGGACGGTAGC 3371
Db 1981 AAAGAGTGTCAACGGGATATCCGTGCATCACAAAAGCGGTTGACACAGGACGGTAGC 2040
Qy 3372 TGATATTAAGAAATATATCCATCAAGAGCTACTCTCAACAATTAATGTGGTCCATTAA 3431
Db 2041 TGATATTAAGAAATATATCCATCAAGAGCTACTCTCAACAATTAATGTGGTCCATTAA 2100
Qy 3432 TCGCCGGTCCCGCACTCGTGTGATTCGTTGACCAAAAGGACAGGGTCAACCTGATCAG 3491
Db 2101 TCGCCGGTCCCGCACTCGTGTGATTCGTTGACCAAAAGGACAGGGTCAACCTGATCAG 2160
Qy 3492 CGGATTCCTATCTTAAGATGAAGGCAAAATCTGTGTGTGATGCGGATCCATACAGCAT 3551
Db 2161 TGGAATCTATCTTAAGATGAAGGCAAAATCTGTGTGTGATGCGGATCCATACAGCAT 2220
Qy 3552 TCCAGGGAAGAAAGTAAAGTTCATGGTCCATTGCCATTAATCCATCAGGTGTATCT 3611
Db 2221 TCCAGGGAAGAAAGTAAAGTTCATGGTCCATTGCCATTAATCCATCAGGTGTATCT 2280
Qy 3612 CGATTTGGGAATACCTTACGATGCGGTAATATGACATTAATCTTGTCAATGTTAGGAC 3671
Db 2281 CGATTTGGGAATACCTTACGATGCGGTAATATGACATTAATCTTGTCAATGTTAGGAC 2340
Qy 3672 CCCGTACAGAAACCATCATCTTACCAACAGTCCAGAGATCAAGCTATCCACACAGCATGCT 3731
Db 2341 CCCGTACAGAAACCATCATCTTACCAACAGTCCAGAGATCAAGCTATCCACACAGCATGCT 2400
Qy 3732 AACGTGAAGGCTGTCCACCACTGAAACATGCGCGAAACATGTGTGCTATAGGTTAG 3791
Db 2401 AACGTGAAGGCTGTCCACCACTGAAACATGCGCGAAACATGTGTGCTATAGGTTAG 2460
Qy 3792 GCTTGTGATGCGGCAACGGAATATCATCATGCGGTTGGCAAGCTCATTAAGGTTTAC 3851
Db 2461 GCTTGTGATGCGGCAACGGAATATCATCATGCGGTTGGCAAGCTCATTAAGGTTTAC 2520
Qy 3852 CCGTGTCTGTAGCTTAAGAAACATGCGGAAATCTGAAGTTCTCTTGTGTTCTTCGG 3911
Db 2521 CCGTGTCTGTAGCTTAAGAAACATGCGGAAATCTGAAGTTCTCTTGTGTTCTTCGG 2580
Qy 3912 CAAGGACAAAGGCAACACACATGACACGAGACACTGGGTATGTGCTTGAACAAT 3971
Db 2581 CAAGGACAAAGGCAACACACATGACACGAGACACTGGGTATGTGCTTGAACAAT 2640
Qy 3972 CTATCAAGGGTCAACAGGTACGAGGAGGAGGAGCTCAAGCTACAGAGTACAGAGG 4031
Db 2641 CTATCAAGGGTCAACAGGTACGAGGAGGAGGAGCTCAAGCTACAGAGTACAGAGG 2700
Qy 4032 TGAATTAAGAAAGCCGCTGACCAAGCTATGTTATGTCTTAATAGCAAAAGGTCAAC 4091
Db 2701 TGAATTAAGAAAGCCGCTGACCAAGCTATGTTATGTCTTAATAGCAAAAGGTCAAC 2760
Qy 4092 AGGTTCCGAGTGTGCGGTGCACTGTACCGAAATGCGCGCTGCTTTTGAATAGACGCC 4151
Db 2761 AGGTTCCGAGTGTGCGGTGCACTGTACCGAAATGCGCGCTGCTTTTGAATAGACGCC 2820
Qy 4152 AATAGCTGTGGGAGCGGCTAGACTTGTGAAGACGAAACCGCTCATCATGCTGTAGG 4211
Db 2821 AATAGCTGTGGGAGCGGCTAGACTTGTGAAGACGAAACCGCTCATCATGCTGTAGG 2880
Qy 4212 ACCCAATTTTCTTAAGATGCGGAAACGAGGAGGAGCCTTAAGCTCGAGTGCCTACAT 4271
Db 2881 ACCCAATTTTCTTAAGATGCGGAAACGAGGAGGAGCCTTAAGCTCGAGTGCCTACAT 2940
Qy 4272 GAGCATAGCGTCCATGCTCAACGCTGACGCGAATTAACAAAAATATCAGTACCGCTACTG 4331
Db 2941 GAGCATAGCGTCCATGCTCAACGCTGACGCGAATTAACAAAAATATCAGTACCGCTACTG 3000

QY	4332	AAACGGCATCTATTCTGGTGGCAAAAGATCGAGTGTGCAATCATGTCACTGTTTAC	43931
Db	3001	AACGTGCATCTATTCTGGTGGCAAAAGATCGAGTGTGCAATCATGTCACTGTTTAC	30606
QY	4392	TGCTTTGCACACTAGCATGCGATGCCATGTGCACATATATTTCTTGATAAACAATGGAGAC	44511
Db	3061	TGCTTTGCACACTAGCATGCGATGCCATGTGCACATATATTTCTTGATAAACAATGGAGAC	31205
QY	4452	CAGATTAATCGAGGCAATTCACCCGAAAGAAAGCTGCAAATTCTG--GATGATGACA	45080
Db	3121	CAGATTAATCGAGGCAATTCACCCGAAAGAAAGCTGCAAATTCTGATGATGATGACA	31805
QY	4509	GCCAGTAGACATYTGACTTGGTCAAGGCTCACCCAAACAGCTCTTTGGCAGGACAGCCAG	45686
Db	3181	GCCAGTAGACATYTGACTTGGTCAAGGCTCACCCAAACAGCTCTTTGGCAGGACAGCCAG	32405
QY	4569	TTACTCCGTCAATGAGGCAAGTTGTATTCATCCTGGAAGGTACACGATTCATCAGAC	46288
Db	3241	TTACTCCGTCAATGAGGCAAGTTGTATTCATCCTGGAAGGTACACGATTCATCAGAC	33000
QY	4629	CGCCAGAGACATTGCGCAAAATTCATGCAATGTGCCCCAACAAATGTGAGCTAATGACA	46888
Db	3301	CGCCAGAGACATTGCGCGAAATTCATGCAATGTGCCCCAACAAATGTGAGCTAATGAGA	33606
QY	4689	GATTTCCTTGACATCCGTTGGGGGAGAGATGTGCAGCATCCGCTCCAAATGCCAGTATGA	47488
Db	3361	GATTTCCTTGACATCCGTTGGGGGAGAGATGTGCAGCATCCGCTCCAAATGCCAGTATGA	34205
QY	4749	GGAGTCAGAGCGTCTGCTCCACCTGCACACACTTCATGCCTGTGTATTAATGCGTATGAC	48080
Db	3421	GGAGTCAGAGCGCTGCTGCTCCACCTGCACACACTTCATGCCTGTGTATTAATGCGTATGAC	34806
QY	4809	GGCTGAGCGCGTATACAGGTTGCCCTCTGCGAAGAAAGACAGTTGCCGTAATGCTATC	48688
Db	3481	GGCTGAGCGCGTATACAGGTTGCCCTCTGCGAAGAAAGACAGTTGCCGTAATGCTATC	35405
QY	4869	ATTCCCTGTTGCCGAAGTACAGATACAGGCGGTGAGAAAGCTACAGTACAGCAACACGT	49288
Db	3541	ATTCCCTGTTGCCGAAGTACAGATACAGGCGGTGAGAAAGCTACAGTACAGCAACACGT	36000
QY	4929	CCTGTTTCAAGCGCTGTACACACCGGCTGTACACCCCGAAGATCGCGGAATTAATCT	49888
Db	3601	CCTGTTTCAAGCGCTGTACACACCGGCTGTACACCCCGAAGATCGCGGAATTAATCT	36606
QY	4989	AGAAACGCGACC 5000	
Db	3661	AGAAACGCGACC 3672	
RESULT 3			
WEEVNS		4498 bp RNA linear	VR1_14-FEB-2003
LOCUS			
DEFINITION	Western equine encephalomyelitis virus partial gene for nonstructural polypeptide, genomic RNA.		
ACCESSION	X74892.1 GI:398206		
VERSION	X74892		
KEYWORDS	nonstructural polypeptide; NSP1; NSP2; NSP3; NSP4.		
SOURCE	Western equine encephalomyelitis virus		
ORGANISM	Western equine encephalomyelitis virus		
REFERENCE	1. Alpha virus; WEEV complex.		
AUTHORS	Yuferev, V. P., Volichkov, V. E., Netesov, S. V., Saifonov, P. F. and Uryvaev, L. V.		
TITLE	Nucleotide sequence of the 26S mRNA and genes of nonstructural proteins NSP2 and NSP3 of Western Encephalitis virus strain 5614		
JOURNAL	Virusologia 28, 86-98 (1992)		
AUTHORS	2 (bases 1 to 4498)		
TITLE	Uryvaev, L. V., Volichkov, V. E., Yuferev, V. P., Samokhvalov, E. I., Lebedev, A. I., Saifonov, P. F. and Netesov, S. V.		
	Primary structure of proteins of the ns2 and ns3 polymerase complex confirm the recombinant nature of western encephalitis virus		

REFERENCE	JOURNAL	Dokl. Akad. Nauk SSSR 335, 813-818 (1994)
AUTHORS	Volchikov, V.E.	
TITLE	Direct Submission	
JOURNAL	Submitted (01-SEP-1993) V.E. Volchikov, All-Union Institute of Molecular Biology of NPO Vector, Koltsovo, Novosibirsk region, 633159, RUSSIA	
FEATURES	Location/Qualifiers	
Source	1. .498	
	/organism="Western equine encephalomyelitis virus"	
	/mol_type="genomic RNA"	
	/strain="5614"	
	/db_xref="taxon:11039"	
	<1. .4498	
	/codon_start=2	
	/transl_except="(pos:4253. .4255,aa:OTHER)	
	/product="nonstructural polypeptide"	
	/protein_id="CA52868.2"	
	/db_xref="gi:28400593"	
	/db_xref="GOA:Q89241"	
	/db_xref="SPTRMBL:Q89241"	
	/translation="TIRKPAVFDSPFVIRLTHSGPDIGLPRSRLLLEPTKPAAPAH TMAVHLRGLQOEAEVAAAEIRALPLPEIKETVEAEVLAQVLMHMGAGAYKE PRGIRVTSYSGEEKISVAILSPQALNSSEKICHPILAEVLMHMGAGAYKE PYHKRVYVEBGTAPVQDFQALSASATVIRNERFVNRYLNLHINSGALNTDEYIK TVKQITDSEYFDIDARKCVKREDAGPLCLTGLVDPPIHEFAVESLTKTPAPAPKV PTIGYGVPSGSGKGIKSAVTKKDLVSAKENCABIRIVRMRMDVARTVDS LNLNKHVPNTLYIDEAPFAGCTLALILAIYKRVKLCDPQCGCFNMMLCKHNF NHDCTEVYHKHSISRCTQVTVTLVSTLFYDKMKTVNCPADKILITDTGTTKHKDD LILCPRGWVKLOIDYKNEHMTAAASOGLTRGVAVYKKNENLYSOTSNNV LVLTTERKIVKTLADGPMIKITPTAKYRGDTFTSLDMQEBHDIMARVLDPKPTADV FONRVNVCMAALEPVLATANIYLTTRQOMETLHFKIDRAVSPMALNFTCTRFQVAD LDSGLFSPATVATLYRDHMDNSGKMGVGLNRAVELSRYPCTIKAVDTGTVADIL RNMTIKDYSPTINVPYLNRLPHSLIYDHKQCTDHSGLSKNGKSVLYIGPISIS MTCXKAEHMLNTGTCVYALIGGLADPANTENITVASFRTPRYCKQKRENTVEYLF VFEKDGNNHTDDRLGVADNITQSTRIYBARAVYRTRKDISKSAQALVNAA NSKQGVSGCAGLYRMPAPAFDRQPLAVGTARLVKHEPLIHAVGNFSPMEDRED LKLAAYMSIASIVNARIRIKISVPLISTGYSGGKRVWQSLHLHTAFPTDADY IYCDKOMETRIEAIHRKESVELDDDDQFVIDLVRVHNSLSLAGPGSVNKGKLY YSYEGTRPHQATLDAIEIHAMMKNSEANOIGLYLIGSMSTRBKCPSESEASA PPHRLPCLCYAMTAERYVYLSRAKKROPAVSSFPKPIRYGVOLQCKPILPSG VPEPAVHPKRYAEILLETTPSPFTTIVCEPTVBERIPSVVISABEASLSLGGVSA FSSASTSSSTAMSDYDRFVYTAADVHANTSTWISIPAPGLVDLPBDTDSHSISPS ALGEVETSPVQDITLACCAAPRGIAEMQFNTEAPFQLSDHREVPAPRRPIPSPS TASAPVPKPRRTYQOPGVARAIISAEIDELVYIROSXNRYEAGAVIYFSETGCHIL QOKVROCKQODPIFERAVHKEYAPRLDLEREMLOKXUQLCASBGNRSRYGSRPI PGLOR"	
mat_peptide	<1. .292	
	/product="NSP1"	
	/product="NSP2"	
	/product="NSP3"	
	/product="NSP4"	
mat_peptide	2675. .4273	
	/product="NSP1"	
	/product="NSP2"	
	/product="NSP3"	
	/product="NSP4"	
mat_peptide	4274. .>4498	
	/product="NSP1"	
	/product="NSP2"	
	/product="NSP3"	
	/product="NSP4"	
BASE COUNT	1279 a 116 c 1125 g 978 t	
ORIGIN		
Query Match	70.7%; Score 3536; DB 14; Length 4498;	
Best Local Similarity	97.9%; Pred. No. 0;	
Matches 3594; Conservative	0; Mismatches 75; Indels 3; Gaps 1;	
1332	AACAATTAGAAATACCTGCGCGCTTGATCATCTTTGTGATTCACAGCCTTACAGCA 1391	
Db	1 AACCAATTAGAAATACCTGCGCGCTTGATCATCTTTGTGATTCACAGCCTTACAGCA 60	
QY	CGGGCTCGATTATGGGCTTCGCGCGTAGGCTCAACACTGCTGTTGAACCAACTGTCAACC 1451	
Db	61 CGGGCTCGATTATAGGCTTCGCGCGTAGGCTCAACACTGCTGTTGAACCAACTGTCAACC 120	
QY	CGACCGGCTATTACAATGCGCGATGTGAGACATCTCGTGAGCTTACAGCAAGAGCTGA 1511	

Db 121 CGCACCGGCTATTTCAAATGGCCGATGTGAGACACTGTGCTTACAGACAGAAAGCTGA 180
Qy 1512 AGAAGTGGCTGCACGCGGAGAGATCAGAGAAAGCCCTGCGACCTTGGCTCCCTGAATAGA 1571
Db 181 AGAAGTGGCTGCACGCGGAGAGATCAGAGAAAGCCCTGCGACCTTGGCTCCCTGAATAGA 240
Qy 1572 AAAAGAGACCGTAGAGCGAGAGATGAGACTTCATTATGCAAGAGCGAGAGCGTAGCGT 1631
Db 241 AAAAGAGACCGTAGAGCGAGAGATGAGACTTCATTATGCAAGAGCGAGAGCGTAGCGT 300
Qy 1632 GGAGACACCGAGAGACATCAGGGTGAACAATTACCAGCGAGAGAGAAAGATTGGCTC 1651
Db 301 GGAGACACCGAGAGACATCAGGGTGAACAATTACCAGCGAGAGAGAAAGATTGGCTC 360
Qy 1692 TTACGCTATCTTCACTCCCGAGGGGTATTGAATAGTAAAGAAAGTGGGTATCCACC 1751
Db 361 TTACGCTATCTTCACTCCCGAGGGGTATTGAATAGTAAAGAAAGTGGGTATCCACC 420
Qy 1752 ATTGGCGGAACAAAGTACTGTATATGACTCACAAGAGTAGGCGAGAGATCAAAAGTGA 1811
Db 421 ATTGGCGGAACAAAGTACTGTATATGACTCACAAGAGTAGGCGGTGGGGAATCAAAAGTGA 480
Qy 1812 GCCATACCAAGTAAAGTCAATTGTACCAAGAGGAGCGGGTCCCTGTTCAAGACTTCA 1871
Db 481 GCCATACCAAGTAAAGTCAATTGTACCAAGAGGAGCGGGTCCCTGTTCAAGACTTCA 540
Qy 1872 GGCAATTGAGTAGAGCGCTACGATCGTTTCAAGAGAGGGAGTTGCTGAACAGATACCT 1931
Db 541 GGCAATTGAGTAGAGCGCTACGATCGTTTCAAGAGAGGGAGTTGCTGAACAGATACCT 600
Qy 1932 GCACACATCGCAATCAACGAGAGCGCTTAAACACTGACAGAGAGTATTAAGACTGT 1991
Db 601 GCACACATCGCAATCAACGAGAGCGCTTAAACACTGACAGAGAGTATTAAGACTGT 660
Qy 1992 AAAAGCTCAGAGACACAGACTGAAATACGTTTGAATTTGACGAGAAAGTGTGTTAA 2051
Db 661 AAAAGCTCAGAGACACAGACTGAAATACGTTTGAATTTGACGAGAAAGTGTGTTAA 720
Qy 2052 GCGGAAAGAGCAGAGTCCCTGTGCTTACCGGTGATGTGATGATCCACCATTTCAAGA 2111
Db 721 GCGGAAAGAGCAGAGTCCCTGTGCTTACCGGTGATGTGATGATCCACCATTTCAAGA 780
Qy 2112 GTTTGCGTACGAGAGTCTCAAGACAGACGACGACGACCTCAAAAGTCCCAACATCG 2171
Db 781 GTTTGCGTACGAGAGTCTCAAGACAGACGACGACGACCTCAAAAGTCCCAACATCG 840
Qy 2172 AGTCTATGAGAGTCCAGGTTAGTAAATCTGGAATCTCAAAAGCGGTGTGACTAAGAA 2231
Db 841 AGTCTATGAGAGTCCAGGTTAGTAAATCTGGAATCTCAAAAGCGGTGTGACTAAGAA 900
Qy 2232 AGATCTGGTTGTGAGTGGAGAAAGAAAGTGGCGCAAAATCTACAGGAGTGTAAAGAG 2291
Db 901 AGATCTGGTTGTGAGTGGAGAAAGAAAGTGGCGCAAAATCTACAGGAGTGTAAAGAG 960
Qy 2292 GATGAGACGTATGAGATGTTGCTGCTAGAGACTGTGATTCAGTGTCTTAAATGGGGTTAA 2351
Db 961 GATGAGACGTATGAGATGTTGCTGCTAGAGACTGTGATTCAGTGTCTTAAATGGGGTTAA 1020
Qy 2352 GCACCCCGTTAAACCCCTGTATGATGAGGCAATTTGCTGTCATGAGGAGAGCTGCT 1080
Db 1021 GCACCCCGTTAAACCCCTGTATGATGAGGCAATTTGCTGTCATGAGGAGAGCTGCT 1080
Qy 2412 GGCACTGATGCGATCGTCAAACTTAAGAAAGTGGTATTTGCGGGGAGCCCAAAACAATG 2471
Db 1081 GGCACTGATGCGATCGTCAAACTTAAGAAAGTGGTATTTGCGGGGAGCCCAAAACAATG 1140
Qy 2472 CGGCTCTTTTAAACATGATGCTGAGAAAGTCAATTTTAAACATGACATATGACTGAGT 2531
Db 1141 CGGCTCTTTTAAACATGATGCTGAGAAAGTCAATTTTAAACATGACATATGACTGAGT 1200
Qy 2532 GTACCATTAAGACATCTCTAGAGAGTGCACACAGACTGTAACCGCATGTCTCCACGCT 2591
Db 1201 GTACCATTAAGACATCTCTAGAGAGTGCACACAGACTGTAACCGCATGTCTCCACGCT 1260

Qy 2592 CTTCTACAGACAGATGAGACGCTTAAACCATGTGCTGATTAATAATCATATAGATAC 2651
Db 1261 CTTCTACAGACAGATGAGACGCTTAAACCATGTGCTGATTAATAATCATATAGATAC 1320
Qy 2652 CACAGGACCAAAAGCCGACAAAGATGATCTGATTTCAACTGTTTCAGAGATGGT 2711
Db 1321 CACAGGACCAAAAGCCGACAAAGATGATCTGATTTCAACTGTTTCAGAGATGGT 1380
Qy 2712 GAAACGCTACATGATGATCAAAAAATCGAAATCATGATGATGCGGCTGATCGCAAG 2771
Db 1381 GAAACGCTACATGATGATCAAAAAATCGAAATCATGATGATGCGGCTGATCGCAAG 1440
Qy 2772 ACTTACCGGAAAGGCGTTTATGCTGACAGTCAAAAGTCAAGAGATCCACTTATC 2831
Db 1441 ACTTACCGGAAAGGCGTTTATGCTGACAGTCAAAAGTCAAGAGATCCACTTATC 1500
Qy 2832 GCAGACTTGTAGACGCTGAACGTTTACTTACACGACAGAAACGCAATTTGTCTGAA 2891
Db 1501 GCAGACTTGTAGACGCTGAACGTTTACTTACACGACAGAAACGCAATTTGTCTGAA 1560
Qy 2892 GAGCTAGCTGTGATCCCTGATTAAGACATTAACGCTTAATATCCGGGGATTTCAC 2951
Db 1561 GAGCTAGCTGTGATCCCTGATTAAGACATTAACGCTTAATATCCGGGGATTTCAC 1620
Qy 2952 GGCTTCAATTGGAAGACTGCGAGCGCAACAGACGCTTATGSCACGCGCTTGTATTA 3011
Db 1621 GGCTTCAATTGGAAGACTGCGAGCGCAACAGACGCTTATGSCACGCGCTTGTATTA 1680
Qy 3012 GCGCGACAGACGCTGATGTGTTCCAAATTAAGTGAACGTTGCTGGCGAAGCTTTAGA 3071
Db 1681 ACCGACAGACGCTGATGTGTTCCAAATTAAGTGAACGTTGCTGGCGAAGCTTTAGA 1740
Qy 3072 GCGAGCTTGGCCACGCGCCAAATGTCGAGAGACAGCAGTGGGAGAGCTTGCAACC 3131
Db 1741 GCGAGCTTGGCCACGCGCCAAATGTCGAGAGACAGCAGTGGGAGAGCTTGCAACC 1800
Qy 3132 ATTCAAGCATGACAGAGCTACTACCTGAATATGSCACTGAATCTTTTTCACAGCTT 3191
Db 1801 ATTCAAGCATGACAGAGCTACTACCTGAATATGSCACTGAATCTTTTTCACAGCTT 1860
Qy 3192 CTTTGGAGTACCTGGACAGTGGGTTATTTCCGCTCTTACCGTGCACCTTACTTACAG 3251
Db 1861 CTTTGGAGTACCTGGACAGTGGGTTATTTCCGCTCTTACCGTGCACCTTACTTACAG 1920
Qy 3252 GATACACACTGGGATTAATCCGACAGGAGAAACATGATGAGCTTAAATGAGAGTATG 3311
Db 1921 GATACACACTGGGATTAATCCGACAGGAGAAACATGATGAGCTTAAATGAGAGTATG 1980
Qy 3312 AAAAGAGTGTCAACGCGATATCCGTGATCAACAAAAGCGGTTGACACAGGAGGTAGC 3371
Db 1981 AAAAGAGTGTCAACGCGATATCCGTGATCAACAAAAGCGGTTGACACAGGAGGTAGC 2040
Qy 3372 TGATATTAAGAAATATATCAATCAAGAGACTCTCCAAATTAATGATGCTTCCATTAA 3431
Db 2041 TGATATTAAGAAATATATCAATCAAGAGACTCTCCAAATTAATGATGCTTCCATTAA 2100
Qy 3432 TCGCGGTTGCGCCACTGCTGATGCTGTAACCAAAAGGACAGGGTCACTGATCAAG 3491
Db 2101 TCGCGGTTGCGCCACTGCTGATGCTGTAACCAAAAGGACAGGGTCACTGATCAAG 2160
Qy 3492 CGGATTCCTATTAAGATGAGAGGCAAAATCTGTTGTTGATGATGCGGATCTTATCAGCAT 3551
Db 2161 TCGGATTCCTATTAAGATGAGAGGCAAAATCTGTTGTTGATGATGCGGATCTTATCAGCAT 2220
Qy 3552 TCCAGGAGAAAGTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 3611
Db 2221 TCCAGGAGAAAGTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2280
Qy 3612 CGATTTGGGAATACCTGACGATGCGGTAATATGACATTAATCTTGTCAATGTTAGAGAC 3671
Db 2281 CGATTTGGGAATACCTGACGATGCGGTAATATGACATTAATCTTGTCAATGTTAGAGAC 2340

QY	3672	CCCCGACAGGAACCATCACTACCAACAGGCGAAGATACCGCTATCCACACAGCATGT	3733
Db	2341	CCCATACAAAGAACATCACTACCAACAGTGGAGATCAAGCTTATCCACACAGCATGT	2400
QY	3732	AACGTGAAGGCTGTCCAACAACCTGAGCACTGGCGGAAATATGTGTGGTATAGGTATGG	3791
Db	2401	AACGTGTAAAGCTGTGCCAACACTGAACACTGGCGGAAATATGTGTGGCATATGGGTATGG	2466
QY	3792	GCTTGCTGATGGCGAACACCGAATATCATCTGCGGTGGCAGCGTCAATTAAGGTTTAC	3855
Db	2461	ACTTGCTGATGGCGAACCGAATATCATCTGCGGTGGTCCCTCAATTAAGGTTTAC	2522
QY	3852	CCGTGTCTGAGGCTTAAGAAACCTGGCGGAAATATGTAGGTCTCTTCGTCTTGGG	3911
Db	2511	CCGTGTCTGAGGCTTAAGAAACCTGGCGAATATGTAGGTCTCTTCGTGTCTTGGG	2580
QY	3912	CAAGGACAAACGGCAACAACAATGACACAGGACAGACTCGGTGTAGTCTTGAACAACAT	3973
Db	2581	CAAGGACAAACGGCAACAACAATGACACAGGACAGACTCGGTGTAGTCTTGAACAACAT	2644
QY	3972	CTATCAAGGGTCCAACACAGGTACGAGGCAAGGAGACCTCCAGGTACAGAGTATCAGAGG	4031
Db	2641	CTACCAAGGGTCCAACACAGGTACGAGGCAAGGAGACCTCCAGGTACAGAGTATCAGAGG	2700
QY	4032	TGACATTAGCAAGAGCGGTGACCAAGCTATCGTTTATGTCTGTATTAACCAAGGTCAAC	4093
Db	2701	TGACATTAGCAAGAGCGGTGACCAAGCTATCGTTTATGTCTGTATTAACCAAGGTCAAC	2766
QY	4092	AGGTTCCGGAATGTGCGGTGCACTGTACCGAAATATGGCCGGCTGCTTTGTATAGACACC	4151
Db	2761	AGGTTCCGGAATGTGCGGTGCACTGTACCGAAATATGGCCGGCTGCTTTGTATAGACACC	2822
QY	4152	AATAGCTGTGGGACGGCTTAACTTTGTGAAGCACGAACCGCTCATCATATGTCTTAGG	4211
Db	2821	AATAGCTGTGGGACGGCTTAACTTTGTGAAGCACGAACCGCTCATCATATGTCTTAGG	2880
QY	4212	ACCCAAATTTTCTAAGATGTCGGGAAACCGGAGGGCGACCTTAAGTCTGAGCTGCTTAAT	4271
Db	2881	ACCCAAATTTTCTAAGATGTCGGGAAACCGGAGGGCGACCTTAAGTCTGAGCTGCTTAAT	2940
QY	4272	GAGCATAGCGTCCATCGTCAACGCTGACGCGATTTCAAAAATATAGTACCGGTCTGTG	4333
Db	2941	GAGCATAGCGTCCATCGTCCAAAGCTGACGCGATTTCTAAAATATAGTACCGGTCTGTG	3000
QY	4332	AACCGGACCTATTTCTGGTGGCAAAAGATCGAGTGAATGCAATCATTTGATCACTGTTTAC	4391
Db	3001	AACGTGACCTATTTCTGGTGGCAAAAGATCGAGTGAATGCAATCATTTGATCACTGTTTAC	3066
QY	4392	TGCTTTGCAACTACGATGCCGATGTCAACCATATATGCTTGGATTAACAATGGGAGAC	4451
Db	3061	TGCTTTGCAACTACGATGCCGATGTCAACCATATATGCTTGGATTAACAATGGGAGAC	3120
QY	4452	CAGGATATGAGGCAATTCACCGCAAGAAAGCGTGAATTTCTG---GATGATGAGAA	4508
Db	3121	CAGGATATGAGGCAATTCACCGCAAGAAAGCGTGAATATCTGATGATGATGAGAA	3180
QY	4509	GCCAGTAGACATTGACTTGTGAGGTCACCCAAACAGCTCTTGTGGCAGGACAGACAGG	4568
Db	3181	GCCAGTAGACATCGACTTGTGAGGTCACCCAAACAGCTCTTGTGGCAGGACAGACAGG	3246
QY	4569	TTACTCCGTCAATGAGGCAAGTTGTATTCATACCTGGAAGGTACGATTCATCAGAC	4628
Db	3241	TTACTCCGTCAATGAGGCAAGTTGTATTCATACCTGGAAGGTACGATTCATCAGAC	3300
QY	4629	CGCCAAAGGACATTTGCCAAATTCATGACATGTGGCCAAACAATTTGAGGCTTAATGAGACA	4688
Db	3301	CGCCAAAGGACATTTGCCAAATTCATGACATGTGGCCAAACAATTTGAGGCTTAATGAGACA	3360
QY	4689	GATTTGCTTTGATACCTCTGGGGGAGAGTATGTCCAGACATCGGCTTCAAAATGCCAGTATGA	4748
Db	3361	GATTTGCTTTGATACCTCTGGGGGAGAGTATGTCCAGACATCGGCTTCAAAATGCCAGTATGA	3420
QY	4749	GGAGTCAAGGCGTCTGCTCCACTTCAACAACCTTCATGCTGTGTAAATTAAGCTATGAC	4808

Dd	3421	GGAATGACAGGGCGTCTGTGCTCACTTCACACACTTCATGCTGTGTAATTACGCATATAC	3486
Oy	4809	GGCTGAGCGCGTATACAGGTGGCCCTTGCGAAGAAGAACAGTTGCCGTAATGCTCATC	4866
Dd	3481	GGCTGAGCGCGCTATACAGGTGGCCCTTGCGAAGAAGAACAGTTGCCGTAATGCTCATC	3546
Oy	4869	ATTCTCTGTTCGCCAAAGTACAGAGATCACAGGCGGTGCAGAACTACAGTACAGAACCGT	4922
Dd	3541	ATTCCCGGTTCGCCAAAGTACAGAGATCACAGGCGGTGCAGAACTACAGTACAGAACCGT	3600
Oy	4929	CCTGTTTTACAGCGCGTACACCAGCGGCTGTACACCCCGAAGTAGTCGGGSAATAATCT	4988
Dd	3601	CTGTGTTCTACAGCGCGTGTACACCAGCGGCTGTACACCCCGAAGTAGTCGGGSAATAATCT	3660
Oy	4989	AGAAAGCCGACCC 5000	
Dd	3661	AGAAAGCCGACCC 3672	
RESULT 4			
LOCUS	U01034	11678 bp RNA linear VRL 02-MAR-2000	
DEFINITION	U01034	Eastern equine encephalomyelitis virus North American antigenic variety nonstructural polyprotein and structural polyprotein genes, complete cds.	
ACCESSION	U01034		
VERSION	U01034.1	GI:39306	
KEYWORDS	nsp1; nsp2; nsp3; nsp4; capsid; E3; E2; 6K; E1.		
SOURCE	Eastern equine encephalitis virus		
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus; EEV complex.		
REFERENCE	1 (bases 1 to 11678)		
AUTHORS	Weaver,S.C., Hagenbaugh,A., Bellew,L., Netesov,S.V., Voichkov,V.E., Holland,D.J., Clarke,D.K., Goussert,L., Scott,T.W., Trent,D.W. and Chang,G.J.		
TITLE	A comparison of the nucleotide sequences of eastern and western equine encephalomyelitis viruses with those of other alphaviruses and related RNA viruses		
JOURNAL	Virology 197 (1), 375-390 (1993)		
MEDLINE	94025587		
PUBMED	8105605		
REFERENCE	2 (bases 1 to 11678)		
AUTHORS	Weaver,S.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (74-AUG-1993) Scott C. Weaver, Biology, University of California, San Diego, La Jolla, CA, 92037-0116		
FEATURES	Location/Qualifiers		
source	1..11678	/organism="Eastern equine encephalitis virus"	
		/mol_type="genomic RNA"	
		/strain="North American antigenic variety"	
		/isolate="82V213"	
		/db_xref="taxon:11021"	
		/c_lone="PEE14 PEE72 PEE59 PEE700 PEE1900 PEE700"	
	1..46		
	47..7528		
		/note="termination codon at position 5681..5683"	
		/codon_start=1	
		/transl_except="(pos:5681..5683,aa:OTHER)"	
		/product="nonstructural polyprotein"	
		/protein_id="AAC53734.1"	
		/db_xref="GI:393007"	
		/translation="MEKVYVDLADSPVYSIQRCPFHEIAYTOYDNDHANARASHLAKLLEGEVDIDVILDIGSAVRTHSHKHHCIPMKASADPRLRYADLKLRKYLDVKICIAKAADLLTVMSPTDAETSLCMHTDSTGRHGSVAAYDVYAHAAPTISIYYOLDKQRTIYWGPDTPTEPMYKNMGALPYTNWADSVEAFANIIGSSDLHERKSFGKVSIMRRKKLOPTNKFIVSGSTIYTERILIRSMHLPNVFHLKGTSFTGRCNTNTIVSEGVVWRKKTISPGLYGKDVLASTHMBEGFLSKVTDTLRGERVSFPVCYVPA TLDDQMGLIATPVSDADAOKLVGLGNRIIVANGRTORNTNNONYLLPVAAQFSRCREHRADLEDKGLAVRERSLVMGCVAFKTKITSITYRPETQITIKVPVNSFVIP QPVSYGDIIDIGLRRIKMLFDAKAPAIITEADVAMHKGLODEAEVAEVAVALP	
5'UTR			
CDS			

```

PLLPEDKTEVADIDILMOBAGASVETPRRIKITTYPGEEMISGSAVSPQAVLN
SEKLAICHPILAEQVLTWMTKGRAGRYKPEPHGVIPDSGAIIPDFOLSEBATTY
FNEBEFYRNYLHHTAANGALNDEEYKYKVSQETSEVDFIDAKKCYKGGAGM
CLUGELVDPREHFAVESLKTTPAPAPKHPVPTIGYVGGSGKSIITSAATKRLUVIS
AKGNCHEIITKDVGRMGMDIARTVDSVLNKGVSVDLTIDBAPFACAGTILALAI
AIVAPKRVVLCDGRQCGFNMCLKVHFNEICTEVYHNSISRCRTVYSITSLIF
YDKMRVTPNPNCKDIITDTSTTKPLKMDIILTCFRGMVQOLQIDYRHEIMETAAQ
GLTRGVYAVRYKVNENPLVAOTSEHNVLLTFRKRIWMTLAGDPMIITLTPASYG
NETATLEMOEHAIDMAKILETPASSDVFNKNVCMKALBEPIDANTILTRSGOM
ETIDAFKDDKAVSPEMALNFCRTCPGVDDISGIFSAPIPLATYTNHNMNSPEPNY
GLCMRTAKELARRIPCLIKAVDTGRVADVRDITKDNPLINVPILNRPHSLVYTH
RYTNGNDYSOLVPMKGTGTVLVGTIPMNIPEKRVETLGPSPQCTYKALDLDGIPALG
KYDIFINVRPYRHHHYOQEDHAIHSMILTRAVHNLNGGICILAGTADARAE
NIISAVARFREFSVCOPKCAMENTEVAFAFGDNKNHLDODRSLVLANIIOGST
OHEGRAPAYRVVGDITKSNDEYIVNAAANKGOGGGVGCALRKRPKPAFDQPVAT
GKALVHSHPRVIVHANGPNRSLENGDOLGSVWMDIARIINNEFTKYSIPILST
GITAAGKORWMSLNHLFTAMDITDADITTYCLDKWESRIKENITRESVEELTEBD
RPVDIELVRVHPLSLAGRPYSTBEKVSYLGTGRFHQADIAIYAMPEKQBA
NEOICLYLVGSMNSIRSKCPVESEASPHPTIPCLCNMTAERYRLRMANKBOF
AVCSSFOLPKRITGVOKIOCSKVPISGTVPPAIPHRKPAVSVEEDTPVYPERLVP
RRAPVPVPVAPRIPSPCTSNGSTSIOSIGEDOSAASAGATISVDHISIMISPSA
TGPRVTRSSSLSEQTPPTPVVYAEIHAOSGLMSIPIITGSTRPSPSPDSRSDITK
TPSASGSHSVDLITPDSVAILEDFRSRSPQFUSEIKPIAPERTVNMNSRSDITK
PIPRKCOVKYTOPPGVAPARISABEDFVRHNRNRYEGAVIFSEEGQHLQOK
STROCKLQYPLERSVHEKFAAPRLDERELQKQOLCSBENRERYOSRKYENK
AIVERLLOIGISYLSAEPQVECYKTVAPMYSSTANSFSAEVAVVCNVLQDE
NEPTVASYNITDEYDAILDMWDGASCLDTPATPCALRSGRSTLRPEIRAVS
PIQNTLONVLAATKRNQNTYOMRELYVLBSAAFNVCCKKYACNDEYMPFKTNPIR
LITENVOYVTKLGPRAALFATKHLNOLHEIPMRPVNLDKVDKVPYGTATHE
RPRKQVIOADPLATAYLCGIHRELVRNLAVLIPMRFDMSABEDFIIAHPOF
GDALLETDIASFDSKEDDAIAMSALMLEDGVDQALNLIEAFNGITVGVNLTGR
FKFGAMKSGMFLTIFNTVYVIMIASRLERLTPSCAFIDNDIVGVNLTDMR
AERGATLNMMEVKIIDAIVGVKARYPGGFIIVGVDIOTIGCRVADPLKPLFKGKPLP
LDDQDQDRRALHDEAARNRIGITBELKAVASRYEAVTVSLITALLTTLASVSIN
FKHTRHPITLYG-
47.1642
/product="nsp1"
1643.4024
/product="nsp2"
4025.5701
/product="nsp3"
5702.7525
/product="nsp4"
7534.11678
7592.11320
/codon_start=1
/product="structural polypeptide"
/db_xref="GI:393008"
/translation="MPPYPLTVPMAPINPMAYRDPNPPRRMRPRPPLAAQIEDL
RSTIANLTKORANPPAGPAPKPKKAPKPAOTKKKPPAPKPKKPKKPKKPKK
RMCKLESDETFPIMANGVNGVACVGGAFRFLHVEGRIDNEQLAAILTKKASITD
LEIGDVQCMKSDTLQYTSDEKPGFTYMHHDVAOVYENNRFTVPGVGGKSGDITLD
NKGVVAVILGCVNBSGRJALSVYTNQKGTADVDEPPSGPMLAVNCLVNIITP
CDQPCMPCCCEKPHETLIMLEQNYDSRAVDOLDAVKNCAARTKDDTHFTQYK
LARVYIADPCPCGHSRCDSPAIIEVNGDAVGIYIOTSMPGLCTKDGVDAWMSM
NGTKOKSIXIDNLHRTSAPCSIVSHHGVIYIACGPDPTVYFHDGPNRHCTVAK
KVERPRVGREKIRHPRPEHVELEPCNRTTHKADQGHYVNAQPELVADHSILSHSK
VKITVPSGAQVKRYCKCPDVAEIGITSSDHTTCTDGVACRAVYLLDNKKNWYSRLBR
GEGDTPFKGLHVPFVYKACIATLAPBEPVLEHGRITLILHLDHDTLTTTRSGD
ANPRQWIERPTTVNFTVTEGLEYLWEPNPKVMAOSEGSEGHMPHEVVVYVYV
RYPLTIIIGCTCAIIMVSCVTSVMLIFTRNLCTIPYKLAPOAVPILALICAK
RPAADDTLOULYNTLANNONNPFMMOTLIPLALIVCMRMLCFCGCPAPILALGAG
AAAYERTAVPNKIGIPIYKALVERPGTALVLOLVNTRIISTENLEITCKKITY
PSPVVCCGATQCTSKPHPDYQCVFTGVVFPFMGAVCFEDENTOMSEYVERSE
CSIDHAKYKHTQVQAMVNIYGSVWSRADVYVNGEPTAKIGDAKLIGPLSSAM
SPDNKYVVYGVYVNDPFEYGTGKASFGDLSRSTNDLANNLKIORQAOI
VHPFQAPBSGEFERKRDKGAPLNDVAPGCSIALPRLAENCAGVSIPTISIDPDA
FTRISETPVUSDIECKITECTYASDREGIATVAKSKSAGNCPIHSSGJAVIYENY
TLBSSGFTTHFSTANIHAPKLOVCTSAITCKGCKCPKHIVIDYPAQHTBSTSAI
SATMSWLKVLVGGTSAFIVGLIATAVVALVLFPHH"
7592.8374
/product="capid"

```

```

mat_peptide 8375..8563
/product="E3"
mat_peptide 8564..9823
/product="E2"
mat_peptide 9824..9994
/product="E6"
mat_peptide 9995..11317
/product="E1"
3'UTR 11321..11678
BASE COUNT 3314 a 2894 c 2809 g 2661 t
ORIGIN
Query Match 55.3%; Score 2766; DB 14; Length 11678;
Beet Local Similarity 72.6%; Pred. No. 0;
Matches 3605; Conservative 0; Mismatches 1355; Indels 6; Gaps 2;
14 ATGATTCATATGGAAGAATTACGTTGATGCTGACGCCGTATGTCAAGT 73
36 ACCATCCAAATGAGAAAGTTCATGTTACCTTAGACGACAGCCATTGCTAAGT 95
74 CGTTACGCGGAGCGTTTCCAAATTTGATGAGAGCAAGGAGGTCATGACAAATGACC 133
96 CACTGCAAAATGCTTTTCCAAATTTGATGAGAGCAAGGAGGTCATGACAAATGACC 155
134 ATGCAATGCGCAGAGCGTTTTCGATGTGCAACAAAGCTCATTTGAGAGCAAGTCAGACC 193
156 ATGTATATGCTAGGGCGTTTTCGACCTAGCTCATGATGAGGAGAAAGTGATATA 215
194 GGGACCAAGTTATCTTGGACATTGGAAGTGGCCCGTCAGACATGCATTTCCATTAACC 253
216 CAGACCGGATGATCTCGAATATTTGGAGCGCGCTGTMAAGCACACCATTCGAAACATA 275
254 GCTATCTGTATCTGCGCCTATGATGATAGGCGTGAAGCCGCGACAGACTCAACGGTATG 313
276 AGTACCTGATATTTGCCAATGATGAGAGCGGACAGACCTTGACAGCTTACCGCTAAG 335
314 CAGAAAGCTTAAAGAAAGTACATTACCGACAAAGACATAGCTTTAAGCGGACAGACC 373
336 CAGACAGCTTAAAGAAAGTACATTACCGACAAAGTATGCTTAAAGCGGACAGACC 395
374 TGGTGAAGTCAATGTCAACACGACGACGAGACTCCATCTCTGTGTATGACACAGACC 433
396 TGCCTAACAGTAAATGTCAACCGCTGACCGCTGAGACACCTGTATGATGATCACACTGACT 455
434 CCAAGTGTAGTACTTTGGAAGTGTAGCATTATACCAAGATGTGTAGCGAGTCATGAC 493
456 CAATTCAGAGTACACAGCGCTCGGCGCTGATATGAGATGTATGAGTGTGATGAC 515
494 CGACATCAATCTACCAACAGCGCTTAAAGAGTTAGACAAATTTACTGATAGGCTTTG 553
516 GCACTTCATTTACTACCAAGCGGTAAAGTGTAGCAATCTATTTGATCGGGTTG 575
554 ACACGACCCCTTTATGTATCAAAAACATGGCAGTTCTTACCTTACTTACACGAACT 613
576 ATACTCAACCGTTCACTATTAAGAACATGACAGCGCTTACCTTACTTACACCAAT 635
614 GGGGTGACGAGAGATTTGGAAGCACGTAACATTTGGCCCTGGTAATCAGATCTTCAAG 673
636 GGGCGGATGAAGTGTGTGGAAGCCGAATATAGGGCTGGGGATGTTGACCTTGACG 695
674 AGACGACGCTTGAAGAACTGCAATCTTGAAGAAAGGCTTCAACTTATTAAGA 733
696 AAAAGATTTGGAAGAAATTCATTAAGAAAGAAATTAACCAACCAATTAAG 755
734 TCATATTTCTGGTTGTTCAACATCTTACACAGAAATAGATCATCTTACGTAAGCTGGC 793
756 TAAATATTTCTGGGGCTCACTATTAATCTGAAGAGAAATATCTTTACGCGAGTTGGC 815
794 ATCTTCAAAACGTTTCACTTGAAGAAAGCTTACCTTACAGAGTAAGTGGGACCA 853
816 ATCTACCTAATGTTTTCATCTAAAGGTAACCTTACCTTACAGAGTAAGTACCA 875
854 TTGTCACTGTGAAGGGTACGTCATCAAAAGATTAACGATCAGCCAGACCTATACGTA 913

```

mat_peptide

Db 876 TCGTAGCTGCGAAGGTTACGTTGTCAGAGAAATACGCTCAGTCTGGGATTTACGGGA 935
Qy 914 AAGTGAAGACTTGGCGTCCACATGCAATCGGAGGGTTCTTGAGTTGCAAAAGTCACAG 973
Db 936 AAGGTATATCTTCTTCGACCATGACCGAGAGGATTTTAAGTTGCAAGGTTACG 995
Qy 974 ATACCTGGCGCGGACGAGAGGGTTCTTTTGCTGTGTGTACGTATGATACAGCACACTT 1033
Db 996 ACACCTTAAGAGGGAGAGGGTCTCTTTCCGTGTGTACGTATGATACAGCACACTT 1055
Qy 1034 GCGATCAATGACAGGAGATTCTGGCAACTGACGTTAGTGTGATGACGCAAAAACTAT 1093
Db 1056 GCGACAGATGACCGGGATACCTGCGACCTGACGTAGTGTGATGACGCGCCAGAAAGCTGC 1115
Qy 1094 TGGTGGGCTCAACCAAGGATTTGTCTGTCATAGTATGAGACCGAAAGAAATACACAA 1153
Db 1116 TGGTTGGCTCAACCAAGGATTTGTCTGTCATAGTATGAGACCGAAAGAAATACCA 1175
Qy 1154 TGCAGAACTATCTTATTCAGAGTGTGCGCCAGCGCTTTCCAGTGGCGCGTGAACATC 1213
Db 1176 TGCAGAACTATCTTATTCAGAGTGTGCGCCAGCGCTTCTCGGGTG---CGGGAAACCC 1232
Qy 1214 GTGCCGACTTGGACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTTACTATGGCT 1273
Db 1233 GCGCAGACCTGGAGGACGAAAGGGGCTAGGGGTACGGGAAGCTTCCCTAGTCATGGCT 1292
Qy 1274 GGTGTGGGCTTTCAAGACCCGAAATACATCCATCTCAAGAAAGCTGTACGCAAA 1333
Db 1293 GGTGTGGGCTTTCAAGACCCGAAATACATCCATCTCAAGAAAGCTGTACGCAAA 1352
Qy 1334 CAATTAAGAAAGTACCTGCGCTCTTGACTCATTTGTGATTCACGCGCTTACAGCGAG 1393
Db 1353 CTATTAAGAAAGTACCTGCGCTCTTGACTCATTTGTGATTCACGCAACCAACAGCTATG 1412
Qy 1394 GGCTGATATGAGGCTTCCGCGCTAGGCTCAAGCTGCTTGAACCACTGTCAAAACCG 1453
Db 1413 GGCTGATATGAGGCTTCCGCGCTAGGCTCAAGCTGCTTGAACCAACGAGACCGG 1472
Qy 1454 CACGGGCTATTCATATGCGGCTATGTGAGCATCTGCGTGGCTTTACAGCAAGAGCTGAAG 1513
Db 1473 CCGCAATTAATTAAGTGAAGCGGAGCTGAGCACTTAAAGGCTGAGAGTGAAGCTGAAG 1532
Qy 1514 AAGTGGCTGACGCGAAGATCAGAGAAAGCCCTGCGCACTTGTCTCCCTGAATAGAAA 1573
Db 1533 CCGTGGCTGAGGCTGAAGCCGTAAGCTGACAGCACTTCTCACTTCTGCGGAGGTGACA 1592
Qy 1574 AAGAGACGCTAGAGCAGAAAGTGAACCTCATTTATGCAAGAGGCGAGAGAGGTAGCGTGG 1633
Db 1593 AGGAGACGCTAGAGCCTGACATCGACCTGATCATGAGAGGCGAGAGAGGACAGCTGG 1652
Qy 1634 AGACACACGAGAGACATCAGGCTGACAGATTACCCAGGCGAAGAGAAAGATTGGCTCTT 1693
Db 1653 AGACACCTTAGACAGACATCAAGGCTGACAGCTACCCAGAGAAAGAAATGATCGGCTGT 1712
Qy 1694 AGGCTATCTTACCCAGGCGGATTAAGTAAAGAAAGAACTGGCGTATCCACCAT 1753
Db 1713 AGGCTATCTTACCCAGGCGGATTAAGTAAAGAAAGAACTGGCTATCCACCAT 1772
Qy 1754 TGGCGGAAACAATCTGTTATGACTCAAAAGTAGGGCAGAGAGATCAAAAGTCGAGC 1813
Db 1773 TAGCTGACAGAGTGTCTGTGATGACTCAAAAGGCGCGAGAGATCAAAAGTGAAGC 1832
Qy 1814 CATACACGCTAGAGTCAATGTATACAGAAAGGAGCGGCTCTGTTCAGACTTCACAG 1873
Db 1833 CATACACGCTAGAGTCAATGTATACAGAAAGGAGCGCTCAATCCCGGATTTCCAGG 1892
Qy 1874 CATTAGTGAAGCGCTGATGCTTTCAAGAGAGGAGTTGCTAAACAGATACCTGC 1933
Db 1893 CTCTAGTGAAGTGAACATAGTATTTAAGAAAGGAGTTGCTTAACCTTACCTTAC 1952
Qy 1934 ACCACATGCAATCAAGGAGGAGCTTAACATGACGAAAGTACTAATAGACTGTAA 1993

Db 1953 ACCACATTCGCTTAAACGAGGAGGATTAATACAGATGAAGAGTACTAACAAGTTGTGA 2012
Qy 1994 AGACTCAGAGACAGACTCAGAAATACGCTTCTGATTAATGACGACAGAAAGTGTAAAGC 2053
Db 2013 AAGCATTGAGACAGACTCAGAAATACGCTTCTGATTAATGACGACAGAAAGTGTAAAGC 2072
Qy 2054 GAGAAAGCAGAGTCTCTGTGCTTAAACCGGATCTGATGATCAGACATTTACAGAGT 2113
Db 2073 AAGGAGATGCGGAGACCAATGTGCTGTGCGGAGACTAGTGAACCCGCACTTCCAGAT 2132
Qy 2114 TTGCGTACGAGAGTCAAGACAGACGACGACCTTCAAAAGTCTTCAACCATCGAG 2173
Db 2133 TTGCGTACGAGAGTCAAGACAGAGTCTTCAAAAGTCTTCAACCATCGAG 2192
Qy 2174 TCTATGAGTGGCAGGCTCAGGTTAACTGTGAATCATCAAAAGGCTGTACTAAGAAAG 2233
Db 2193 TTTATGAGTGGCAGGCTCAGGTTAACTGTGAATCATCAAAAGGCTGTACTAAGAAAG 2252
Qy 2234 ATCTGTTGTGAGTGCAGAAAGGAAAGAACTGCGCAAAATCATCAGGATGTAAGAGGA 2293
Db 2253 ATCTGTTGTGAGTGCAGAAAGGAAAGAAATTCGATGAAATCATTAAGCTCAAAAGTGA 2312
Qy 2294 TGAAGCTATGAGTGTCTGCTAGAGCTGTGATTCAGTGTCTTAAATGGGTTAAGC 2353
Db 2313 TGCGCGCATGAGACATCGCCCGCAGAGTGAATTCGCTACTGCTAATAGGGGTAAAC 2372
Qy 2354 ACCCGTTAAGCTCTGATCAATGATGAGGATTTGCTGCAACGAGAGCGCTGTG 2413
Db 2373 ACTCGTGACACATGATCAATGAGGAGATTCGCTTGCATGACAGAGACCTGTGAG 2432
Qy 2414 CACTGATTCGATGCTCAAACTTAAAGAGTGTATTTGCGGGGAGCCAAAGCAATGCG 2473
Db 2433 CACTGATTCGATGCTCAAACTTAAAGAGTGTATTTGCGGGGAGCCAAAGCAATGCG 2492
Qy 2474 GCTTCTTTAAGATGATGCTGAAAGTCAATTTTAAACATGACATATGACTGAAGTGT 2533
Db 2493 GCTTCTTTAAGATGATGCTGAAAGTCAATTTTAAACATGACATATGACTGAAGTGT 2552
Qy 2534 ACCATTAAGATCTCTAGAGTGTGACACAGACTGTAAACCGCATCTGTCTTCAAGCTCT 2593
Db 2553 ATCAACAGATATTTCTGCGGAGTGCATGAGACGATCAATCAATCTTCTTCAACCTGT 2612
Qy 2594 TCTACGACAGAGATGAAGACGCTTAAACCATGCTGATTAATTCATATGATATCA 2653
Db 2613 TCTATGATTAACGATGAGAACTGTCAACCATGATGATTAATTCATATGATATCA 2672
Qy 2654 CAGGACCAAAAGCCGACAAAGATGATCTGATTTCTAACTGTTTACAGAGATGGTGA 2713
Db 2673 COAGTATTAACAAACCTTTAAAGATGATTAATTTAACTGTTTACAGAGATGGTGA 2732
Qy 2714 AACGCTTACAGATTAATTAACAAATCAAGAAATCATGACTGCGGCTGATGCAAGAC 2773
Db 2733 AACGCTTACAGATTAATTAACAAATCAAGAAATCATGACTGCGGCTGATGCAAGAC 2792
Qy 2774 TTAAGCGGAAAGGCTTATGCTGTCAAGTCAAGTCAAGAAATCACTCTGCTGCG 2833
Db 2793 TTAAGCGGAAAGGCTTATGCTGTCAAGTCAAGTCAAGAAATCACTCTGCTGCG 2852
Qy 2834 AGACTTGAACAGTGAACGTTTACTTAACGACGACGAAAGAAAGCAATGCTGGAAGA 2893
Db 2853 AGACTTGAACAGTGAACGTTTACTTAACGACGACGAAAGAAAGCAATGCTGGAAGA 2912
Qy 2894 CGCTAGCTGATGATCTGTGATTAAGACATTAACGCTAAATATCCCGGGATTTACAGG 2953
Db 2913 CTTGGCCGCTGACCTTGTGATCAAGAGCTTGAACGATGATCAAGTATTTCAACCG 2972
Qy 2954 CTTGATGAGAGCTGAGGCGGACGACGACGACGATATGACGAGCGCTTGTGTAAGC 3013
Db 2973 CCACTGAGAGAAATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGC 3032
Qy 3014 GCGACAGCTGATGATGTTCCAGAAATAGGTGAAGCTGTGCGGCGAAAGCTTTAGAGC 3073
Db 3033 CAGCTAGAGGAGCGTTTCCAAATTAAGTGAAGTGTGCTGCGGCGAAAGCGGTAGAGC 3092

QY 3074 CAGTCTGGCCAGGCGCAACATGTTGCTGACGAGACAGTGGGAGAGCTTTGACCCAT 3133
 Db 3093 CTGTGTTGGCCACCGCCCATATTTAGCTGACCCGCTCGCATGTTGGAGACTTTTCAGCGCT 3152
 QY 3134 TCAGACATGACAGACGCTACTACCTGAAATGGACCTGAACTTTTGGACACAGTTCT 3193
 Db 3153 TCAGAGATGACAAAGCGTATTGCGCTGAGATGGCTTAACTTTTTCGACACAGATTCT 3212
 QY 3194 TTGAGATGACCTGACAGTGGTTATTTTCCGCTCTTACCGTGGCACTTACTTACAGGG 3253
 Db 3213 TTGGCGTGCATGACAGCGGGTGTCTCCGCGCCAACTGTTCCGCTGACTTACACCA 3272
 QY 3254 ATCAGACTGGGATTAATCCGCGAGGAAAGACATGTAATGGCTTAATAGAGAGGTGCAA 3313
 Db 3273 ATGAACTGGGATTAATAGCCCAAGGTCCAAACATGTAATGTTTGGACATGGCACTGCTA 3332
 QY 3314 AGGAGTTGTCAGCGCGATATCCGTGATCACAAGAAAGCGTTGACACAGCGAGGTAGCTG 3373
 Db 3333 AAGAACTTGCAGCTCGGTATCTTGTATCTGAAAGCGGTGATACAGGTAGAGTGGCTG 3392
 QY 3374 ATATAGGAATTAATACCTCAAGACTTCTCCAACTTAATATGTTGTTCCATTAAATC 3433
 Db 3393 ACGTTCGACAGACATATCAAGACTATMAACCGCTAATTAATGTTGTTACCCCTGATA 3452
 QY 3434 GCCGATTGCCCATCTCGTTGATCGTTGACCAAGAGACAGGATACAACTGATCAGACGG 3493
 Db 3453 GAAGACTCCCACTCATTTGTTGTCAACATATGATACCTGGAAACGTTATTAATCTCC 3512
 QY 3494 GATTCTATCTATAGATGAGGCAATCTGTTGTTGATCGCGCATCTTACAGCAATC 3553
 Db 3513 AGCTAGTACCAAGATGACCGGAAACCGTACTAGTGGGTAGACATATGAACATAC 3572
 QY 3554 CAGGAAAGAAAGTGAAGTCCATGGTCCATTGGCCCACTAAATCCATCAGGTGTATCTCG 3613
 Db 3573 CAGAAAGAGAGTGCAGACATAGGCCCAAGCCCAATGATACATATMAACGGAACCTGG 3632
 QY 3614 ATTGGAATATCTAGGCATGTCGTTAAATATGACATATCTTTGTCAATTTAGGACCC 3673
 Db 3633 ACTGGGATCTCTGCGCTTAAAGCAATATGACATATTTTATTAACGTAGAGATC 3692
 QY 3674 CGTACAGGAACCATCATCTACCAAGTGCAGAGATCAAGCTATCCACCAACAGATGCTAA 3733
 Db 3693 CTTACCGACATCACCATTTACCAACAGTGCAGAGACATAGCATCAACACAGCATGCTTA 3752
 QY 3734 CGTGAAGCTGTCACTCACTGAAACACTGGCGGAACATGTGTGCTATAGGGATGGGC 3793
 Db 3753 CCAGAAAGCAGTGACCATTTGAAACAAAGCGGTACGTGATCGCATTTGGGCTATGGGA 3812
 QY 3794 TTGCTGATCGGCAACGAGAAATATCATCTGCGGGTGGCAGGCTAATTTAGTTTACC 3853
 Db 3813 CTGCGGACAGGCAACGAGAACTATCTCGGAGTGCAGCCGCTCATTCAGGTTCTAC 3872
 QY 3854 GTGTCTGTCAAGCTTAAGAACTGCGGAAATACTGAGGTTCTTTCGTGTTCTTCCGA 3913
 Db 3873 GTGTGTGCAGCCGAAAGTGTGCTGGGAAACACTGAGGTGCGCTTGTGTTTCCGCA 3932
 QY 3914 AGGACAAACGGAACCAACATGACAGAGCAACTCGGTGTGTGTGCTTACCAACATCT 3973
 Db 3933 AGGACAAACGGAACCATCTCCAAATCAAGATAGAGTGTGTGTTAAACAACATAT 3992
 QY 3974 ATCAAGGATCAACAGGTAGAGGAGGAGAGGCTCCAGCTACAGAGTATCAGAGGTG 4033
 Db 3993 ACCAAGGATCACTCAACATAGAGTGCAGAGCACTGCAATATAGAGTGTGCGCGCG 4052
 QY 4034 ACATTAGCAAGAGCGCTGACCAAGCTATCTTAAATGCTGTCTAATATAGCAAGGTCAAC 4093
 Db 4053 ACATTAACAAAGAGCAATGATAGTTATGTTAAGCGGCAACAAAGGGAACCCG 4112
 QY 4094 GTTCCGAGTGTGCGGTGACATGACCAAAATGGCGGCTGCTTTGATGACAGCCAA 4153
 Db 4113 GTGCGGTGTGTGTGTGTCCTTTTACAGGAAGTGGCTGAGCTTTTGTATGAAGCGCG 4172

QY 4154 TAGCTGCGGAGCGCTAGACTTGTGAAGCAGAACCGCTCATATCATATCATGTGTAGAC 4213
 Db 4173 TAGCACTGGTAAGCCACCTTGTCAAGCATTTCTCCGAACGTATCATGTGTGTGCC 4232
 QY 4214 CCAATTTTCTTAAGATCCGGAACCGGAGGCGCACTTAAGCTGCGAGTGTCTTACATGA 4273
 Db 4233 CTAAATTTTCTCGCTATCAGAAAACGAGAGAGCCGAAATTTGTGGAAGTATACATGG 4292
 QY 4274 GCATAGGTCATATCGTCAAGCTGAGCGGATTAACAAAATATAGTACCGCTACTGTCAA 4333
 Db 4293 ACATTCGCAAAATTAATACAAACGAGAGTTTACTTAAAGTCTCATTCCTGTTATCTA 4352
 QY 4334 CCGGCATCTAATCTGTGGCAAGATGAGTATGATGATGATGATGATGATGATGATGATG 4393
 Db 4353 CCGGCATCTAAGCAGAGTGTATAGAGACAGGTTATGCAATGCTGAAACATCTATTTACAG 4412
 QY 4394 CTTTGAACCTAGCGATGCGGATGTCACATATATGCTTGTGATTAACAAATGGAGACCA 4453
 Db 4413 CCATGATTAATACGAGCAGACATTAATTAATGCTGTATGATATACAAATGGAGTCAAA 4472
 QY 4454 GGATTAATGAGCGCATTCACCGCAAGAAAGCGTCCAA---ATTCTGATGATGACAAAGC 4510
 Db 4473 GAATTAAGAAAGCTATCATCTGGAAGAAAGCGTTGAAGCTTACTGAGATGACAGAC 4532
 QY 4511 CAGTACACATTTGACTTGTGTCAGGTCACCCCAACAGCTCTTTGGACGAGACCAAGTT 4570
 Db 4533 CAGTGAACATTTGAATCTGTAGGCTGACCGTTGAGCACTTGGACGATGACCTGTGTT 4592
 QY 4571 ACTCGTCAATGAGGCAAGTGTATCATATCTGGAAGTACAGATTCATTCAGACCG 4630
 Db 4593 ATTCAACACCGAGGCGAAGGTGATCTGATCAGAGGAGTATGATTCATCAACATCG 4652
 QY 4631 CCAAGCAATTTGCGGAATTCATGCAATGAGGCGCAACAAATGAGGCTTAATGAGCAGA 4690
 Db 4653 CCAAGCAATGAGTGAATTTAGCTATGAGGCTTAACAGAAAGCAAGAAAGCAGCAGA 4712
 QY 4691 TTGCTGTATCATCTGTGGGAGAGATGATCTCAGACATCCGCTCAATGCGCCAGTAGAGG 4750
 Db 4713 TTTCCTTATAGTGTGTGGAGAGATGATGAACAGCATCCGCTTAAGTGTCCAGTTGAAG 4772
 QY 4751 AGTCAGAGCGTGTCTCACTCACTCAACATTCATGCTGTGTATTAATGCTATGACGG 4810
 Db 4773 AGTCGAGGCGCTCTTCCCGCTCAACATCCGCTGTGTGCTCACTATGACATGACCTG 4832
 QY 4811 CTGAGCGCGTATCAGGTTGCGCTGCGGAAAGAAAGCAAGTCCGCGTATGCTCATCAT 4870
 Db 4833 CAGAGCAGTGTACAGATTACATGATGCAAGAAATGAAACATTTGCGTGTGCT 4892
 QY 4871 TCTGTGCGGAAGTACAGATGATCAGAGCGCTGCAAGAACTACAGTGCAGCAACAGTCC 4930
 Db 4893 TTCACTTACCGAAATACAGATTAACAGGGGTTGAGAAATTTCAATGCAATGAACCTGTGA 4952
 QY 4931 TGTTTCAGCGGTGTACCAACCGGCTGTACACCCAGGAAGTACCC 4976
 Db 4953 TATCTCTGCACTGTACCCCGCCATCATCATCAAGAAATTCGC 4998

RESULT 5
 BEEVIRNA
 LOCUS 11675 bp RNA linear VRL 08-JUL-1992
 DEFINITION Eastern Equine Encephalomyelitis virus RNA sequence.
 ACCESSION X63135 X67111
 VERSION X63135.1 GI:59185
 KEYWORDS 6K protein; C protein; E1; E2; E3; NS1; NS2; NS3; NS4.
 SOURCE Eastern equine encephalitis virus
 ORGANISM Eastern equine encephalitis virus
 Viruses; ssRNA positive-strand viruses; no DNA stage; Togaviridae;
 Alphavirus; BEEV complex.
 1 (bases 1 to 11675)
 REFERENCE
 1 Volchkov,V.E., Volchkova,V.A. and Neteov,S.V.
 Complete nucleotide sequence of the Eastern equine
 encephalomyelitis virus genome
 Mol. Gen. Mikrobiol. Virusol. 5, 8-15 (1991)

MEDLINE 91375524
PUBMED 1896061
REFERENCE 2 (bases 1 to 11675)
AUTHORS Volchokov, V.E.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-1991) V.E. Volchokov, All Union Institute of Molecular Biology, Koltoovo, Novosibirsk Region, 633159, USSR

FEATURES
source
1. 11675
/organism="Eastern equine encephalitis virus"
/mol_type="genomic RNA"
/strain="ssp. North American variant"
/db_xref="taxon:11021"
/cell_line="virus grown on vero cells"
/clone_11b="cDNA plasmid clones grown in E.coli JC-5183"
47. 5683
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA44846.1"
/db_xref="GI:59187"
/translation="MEKVHVDLSDSPFVSKLQRCFPFEIEATQVTDNDANRARS
HLATKLEGEVDITDOVILIDISAPVRHSHKTHCICPMKASBPDLVYADLRK
SDVDKCIASAAADLLTWSTPDAETPSLCMTDSTCKHSVAVYQDVAVHAPS1
YQALKGVRITVWIGFDTTPMYKMAAGAYPTYNMADESVELEARNIGLSDLHEK
SFGKYSIMRKKKLOPTNKVIFSVGSTIYEEIRLLRSMHLPNVFLHKGTSFGRCNT
IVSCGGVYKKITLSPGIVGVNLSATMRHGFSLCKVDTLGERYSFVCTVYPA
TLCDMGITLADTVSDVDAOKLIVNORITVNGRTONTNTMONTLLPYVQAFCRC
REYRADLEDEKGLGRERSLVWGCGMAKTHKITSTYKRGQTITKAVAFNSFVLP
QPTVGPDIGLRRIKMLFDAKKAAPLITADVAHLKGLDEAEVAEVAEVAALP
PLPEVDEKETEADIDLMOAGASVETPRRHIVTSPGEIMIGSAVAPVALN
SEKLACHLAEQVLMTHKGRAGRYVEPRHRIKVTYVPEEMIGSAVAPVALN
FNEBEFVRVYHLHIAVNGALNTDEEYKVKVSTETDESEVFDIADKKCKGGLAPV
CLAEIVDPREHFAVESLKTBRPAAPKVPITGVYGVSGSGKGIKSAATYKRLVYS
AKKNCHEITIDVKRMGMDLAATVDSVLNAGVSHVDITLIDBAFCHAGTLLALI
AIVKPKVNLCDPQKQCFNMKCLKVFHEHLETCVEVHKSISRCKTVTSVLE
YDKMRIVNPNCDKIIDTSTYKPKODIILTCRGWVKQLQIDYKHEIMTAASQ
GFTKGVAVAYKVENEDLVAOTSSEHVVLLRTEKRIKMTKLADPMIKTLTASVPG
NFTALTEMOAEHDAIMAKILETPASSDVFGNKVAVAKALEPILATANTLIRSOW
ETIPAKODKAVSEPMALNFCRTPRGVIDDGLPSAFTPLVSTNHEMDSRPNMY
GLCHRNAKELARRPCILKAVDTGRVVDRTDKINPLINVPFLNRDLHSLVAVTO
RYTNGDYSQLVTKMTKTVLVGTPMNI PKRVEVTEKINPLINVPFLNRDLHSLVAVTO
KYDII FINVTRPYRHHYQOCEDEHAIHSMLTRAVDLNKGCTICLAGXTAARATE
NII SAVASFRFSYCCPKCAMENTEVAFFVGONGHLDODRSLVANNIOGAT
OHEAGRAVAVVRGDIITKSNDEVIVNANKKGGPCCGALATKMKVAGADKOPVAT
GKALVHGS PNVIVAVGNPFRSLSENEGDQKLSRYNDIAIINNERTKXISILST
GIYAGGKDRVQSLNHLFTAMDITDADITITCLDQWESRIKEAITRESEVEELTED
RPVIDELVRHPLSLAGRPYSTEGRYVYLRTGRHQAKOVAELINAMPMKOEAF
NEOICLVLGESMNSIRSKPVESEASPPHITPCLCNVMTAERVYRLMAKNOEF
AVGSSFOLEPKRITGVOKIQCSPVIFSGTVPAPAIHPKFAVTEDEPVVQERLVP
RRPAPVPVPAIRISPPCTSTNGSTISIOSLGEYQASAAAGCAETISVNDIPRPA
TGFVTRSSSLSEOPTPTPVVVEAETHASGSLMSITISITGFEFRVDSPPSODRS
TPSASASTFVDLITPDSVAILEDPSFSPQFLSEKPIPAFTRVNNMSRSADTJK
PIPKRCQVXYTQPPGVARISAEPDEFVRHSN"
47. 1642
/product="NS1"
1643. 4024
/product="NS2"
4025. 5680
/product="NS3"
5681. 5683
/note="stop codon between NS3 and NS4 genes"
5702. 7528
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA44847.1"
/db_xref="GI:1334981"
/translation="YIRSSSEGGHLOOKSROCKQYPILESVHEKFAVPLDLER
EKLQKQLQCLASBGNRSYOSRKVENKALITVEKLIGISYSIAEQPECKKVY
PAPMYSSTASNSFSSAEVAVCNVLVDENPTVAVYNIIDEYADYLDMDVAGCCID
TATCPAKIRSFPKKHSYLRBEIRSAVSPIONTLQDVLAATATRNCAVTRRELPLV

DSAAFNVECFKKYACNDYWDYFVNTNIRLTAENVYQVTKLAKPKAAALFAKTHNLO
PLHEIPMDRMDKRDVKTPTGKTEERKQVIOADADLATAVLCGIRBELVRL
NAVLINHTLFEDKASADPDALIAEHPOFGAVLETDIASPDKSDEADIANASALMIE
DLGVQDLNLIIEAARGNITSYHLPTCRPFKGSMSGMLTPIFIVNINIMASRV
LREHITTSPCAAPFGDNTYKGVTSDLMEERCATLMMEKVIIDAVGVKAPYFCGG
FIVVDQITGCAVRADPLKRLFYKGLPLDDDDVRRRLAHDEAARMNIGITEEL
VKAVESRYENVVYSLIITALLTASSVNFHGRHGHITLVG"
5702. 7525
/product="NS4"
7526. 7528
/note="stop codon between unstructural and structural
genes"
7534
/note="start for 26S sub-RNA"
7592. 11317
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA44845.1"
/db_xref="GI:59186"
/translation="MPPEYPTLVPPMAPIINMAVRDPPRRRRPPRPLAQTEDL
RRSIANTTLKORANPPAGPPAKKKAAPSLISLRKKRPPPPAKKQRRKRRGROR
MCMLESDKTPEIMLNGQVNGYACVVGARFEPKLVHVGRLDNEQLAIKLKAS1YDL
EYGVQPCMSDITQYTSDBRPGFVNMHGAQYENNRFTVPRGVGGKSGSRILDN
KGRVVALVLGVEGSRALSVTVMNKGVTYKOTPGSEBMSLATVWCIVANITFPC
DOPPCMCCEKQNPHEHTLMLEONYDSRAVDOLLDAVKKCARTRRDLDVHTFOYK
ARVYIADPCNCGSRCSPLAIEVRDADAAGVRIOTSANRFGKTRDLDVAVNSFMN
GRTOKSKIINDLHRTSAPCSLVSHHGYIILAQCPGQVTVGHDEGNHRTCTVAHK
VEFRPVERERYRHPHEGVLPNCRVYHKAADQGVYEMQPLVAASHLISHAKY
KITVPSGAQVYKCPKDPVEGTSSTYTTCTDVKOCRAVLIIDNKVMVNSGLPRG
EGDFEKKLHPVEPVYAKCIATLAPPLVEHKKRIILHLPHPTLLTTRSGSDA
NPQWIERPTVVEFTVGTGLEVTWNHGWKWALESGBGNHGMHVEVYVYNR
VPLITIGCTVCAIIVWSCVTSYKTRNICTPYYKAPNAOVPIIALLCCIRP
TRADDTQVANTLVNMMNPFMMOTLPLIALIYCMMLKLPFCGSAFLVLCALGA
AAETHVMPNKVGI PYKALVERGVA PVHQLQVLNTRILPSTNLVITCKYTKVP
SPVVCAGAOCTCKSPHDOVOCQFTVGFPMWGAFCFDENTOMSEAVESSEB
SIDHAKYKATGQCAQMANIYVGSVMSRADVANEETPAKIDAKLIGPLSAS
PFDKVVAHGEVYVNPPEYGTAKSGSPGKORSTSDNLVNTUKLORPAGIV
HTPEPVAHSGEHRKRGAPINDVAFPGSIALERPARNCVNSIPISIDIPDAF
TRISEPTVSDLEKKTITECTYASPGSIALVAYSSAGCNPISBPGAVIKENDYT
LAEBSGTFHFSHNPAPFLQVCTSAVTCGDKCKPHIVDPQMHTEFSTAIS
ATAMSMKLVGVGTSAPIVGLIATAVALVLPFHRH"
7592. 8371
/product="c protein"
8372. 8560
/product="B3 protein"
8561. 9820
/product="E2 protein"
9821. 9991
/product="6k protein"
9992. 11314
/product="E1 protein"
ORIGIN
BASE COUNT 3322 a 2882 c 2805 g 2666 t

Query Match 54.9%; Score 2745.2; DB 14; Length 11675;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 3592; Conservative 0; Mismatches 1368; Indels 6; Gaps 2;

QY 14 ATGCATCCAAATATGGAAGAATTCACCTTGACTTGAATGTGACAGCCCGATGTCAAGT 73
DB 36 ACCATTCMAAATGAGAAAGTTCATGTGACTTGAACGACGAGACCCCATTCGTCAAGT 95
QY 74 CATTACAGCGGACCTTTCCACAAATTTAGATGGAAGCAGGCGAGTCACTGACAATGACC 133
DB 96 CACTGCAAAAGATGCTTTCACATTTGAGATAGAGCAAGCAGCAGTCACTGACAATGACC 155
QY 134 ATGCATGCGAGGCGCTTTTTCGATGTGCAACAAGCTCATTTGAGAGGAAGTGCAGC 193
DB 156 ATGCTTAATGCTAGGGCGCTTTTCGACCTAGCTAATTAAGCTCATTTGAGAGGAAGTGCAGT 215
QY 194 GGAACCAAGTTATCTTGACATTTGGAAGTGGCCCGCTCAGACATGACATTTCCAAATCACC 253

D	b	216	CAGAC	CAG	GTAT	CCT	TGG	AAT	TGG	GAG	CGG	CGT	GTA	AGG	CAT	ACG	ATT	CCAA	CA	T	275	
O	y	254	GCT	AT	CA	T	T	GT	AT	CT	G	CC	CT	T	AT	AT	A	AG	CC	GG	A	313
D	b	276	AGT	AC	CA	CT	GT	AT	T	T	GG	CC	AA	T	GA	AG	CG	C	AG	AA	335	
O	y	314	CAG	AA	AG	CT	TA	AG	AA	AG	T	GA	CA	T	T	AC	CG	A	CA	AG	373	
D	b	336	CAG	AC	AA	CT	TA	AG	AA	AG	T	GA	CA	T	TA	CG	CA	AA	T	GA	395	
O	y	374	TG	CT	GA	AG	T	CA	T	GA	CA	CA	CA	GG	A	GA	CT	CA	T	CT	433	
D	b	396	TG	CT	TA	CA	GT	AT	AT	GT	CG	AC	CG	CT	GA	CA	CA	CC	CT	GT	455	
O	y	434	CA	CG	CT	GA	GT	AG	T	CT	T	T	GA	AG	T	GA	CA	T	GA	CA	493	
D	b	456	CA	AC	TT	GA	AG	GT	AC	CA	CG	GC	TC	CG	GT	AT	AT	AT	CA	AG	515	
O	y	494	CG	AC	AT	CA	AT	CT	CA	CA	CA	GG	CG	CT	T	AA	AG	A	GT	TA	553	
D	b	516	CG	AC	TT	CA	TT	TA	CT	CA	CA	GG	CG	CT	GA	AG	GT	TC	GA	AT	575	
O	y	554	AC	AG	CA	CC	CC	TT	T	TA	T	GT	TA	CA	AA	AA	CA	T	GA	AG	613	
D	b	576	AT	AC	TA	CA	CC	GT	TA	CT	CA	GT	TA	GA	AA	CA	AG	CG	CG	CT	635	
O	y	614	GG	CG	CT	GA	CG	AG	TA	T	T	GA	AG	CA	CG	T	TA	A	C	T	673	
D	b	636	GG	CG	CG	AT	GA	AA	GT	GT	TT	GA	AG	CA	AA	AT	AT	A	G	GG	695	
O	y	674	AG	AG	CA	GC	CT	T	GA	AA	A	CT	CT	CA	T	CT	T	TA	GA	AG	733	
D	b	696	AA	AA	AG	AT	TC	CG	AA	AG	AT	TC	CA	T	T	AG	CG	AA	AA	AT	755	
O	y	734	TC	AT	TT	CC	GG	T	T	GG	TT	CA	CA	AT	CT	CA	CA	GA	AT	AT	793	
D	b	756	TA	AT	AT	TT	CT	GT	GG	GT	CA	CT	AT	CT	AT	CT	GA	AA	GA	AA	815	
O	y	794	AT	CT	TC	CA	AA	CG	T	TT	CC	AT	TT	GA	AA	AG	AA	AG	T	CT	853	
D	b	816	AT	CT	AC	CT	AA	AT	CT	TT	CA	CT	AA	AA	AG	TA	AA	CT	AT	AG	875	
O	y	854	TT	GT	CA	GT	GT	GA	AG	GG	T	CG	T	CA	T	CA	AA	AA	AG	TA	913	
D	b	876	TC	GT	CA	GT	CG	AA	GG	T	TA	CG	TT	CA	AG	AA	AT	TA	CC	GT	935	
O	y	914	AA	GT	TA	GA	AA	CT	T	T	GG	CC	TC	CA	AT	AG	CA	CG	GA	AG	973	
D	b	936	AA	GT	GA	AT	AT	CT	TT	CT	TC	GA	CA	T	GA	CA	CG	GA	AG	TT	995	
O	y	974	AT	AC	CT	CG	CG	CG	CA	AG	GG	TT	CT	TT	T	G	CT	GA	TC	GA	1033	
D	b	996	AC	AC	CT	TA	AG	AG	GG	A	GG	T	CT	TT	CC	CG	T	GT	TA	CG	1055	
O	y	1034	GG	CA	T	CA	AT	GA	TA	CA	AG	GA	T	T	CG	CA	CT	GA	CT	GA	1093	
D	b	1056	GG	CA	CA	AT	GA	TA	CA	CG	GA	T	ACT	GG	CA	CT	GA	GT	CA	GT	1115	
O	y	1094	TG	GT	T	GG	CG	CT	CA	CA	AA	AG	A	GT	TC	GT	CA	AT	GA	CG	1155	

QY	1334	CAATTTAGAAAGTACCTGCCGCTCTTTGACCTATTTTGATTTCACGGCTTAACGACGAC	1393
Db	1353	CTATCAAGAAAGTCCCGCCGCTATTTAATCTTTGTCATCCCAACCAACAGCTATG	1412
QY	1394	GGCTCGATATAGGGGTTCCCGCGTAGGGCTCAAGCTGCTGTGAACCAACTGTCAAAACCCG	1453
Db	1413	GGCTCGATATAGGATTTGGTTCGCCGAATTAAGATGCTATTTCAGCGAAAGAACACCCG	1472
QY	1454	CACCGGCTATTACAAATGGCCGATGTGGAGCATCTGCGTGCTTACAGCAAGAACTGAG	1513
Db	1473	CTCCAAATATTATCTGAGGGCCGACCTGTGACACTTTAAAGGCTGCGAAGATGAAGCTAAG	1532
QY	1514	AAGTGGCTGCGCGGAAAGATTCAGAGAAAGCCCTGCCACCCCTTGCTCCCTGAATTAGAA	1573
Db	1533	CCGTGCTAGAGCTGAAGCCGTACGTGACGATTAACCTCCACTTCTGCGGAGGTCCGATA	1592
QY	1574	AAGAGACCGTAGAGGCAAGATAGACCTCATTTATGCAAGAGCAGAGCAGTGAAGCGTG	1633
Db	1593	AGGAGACCGTAGAGGCTGACATTCACCTGATCATGCAAGAGCAGAGCAGGACGCTGG	1652
QY	1634	AGACACCAAGAGACACATTCAGGGTGACAACTTACCCAGGCGAAGAGAATTTGGGCTT	1693
Db	1653	AGACACCTAGACGACACATCAAGGTCAAGATTCAGAGGGAAGAAATGATCGCTGT	1712
QY	1694	ACGCTATCTCTTCAACCCGAGCGGTATTTGAATAGTGAATAAATGGCGGTATTCACCCAT	1753
Db	1713	ACGAGTCTCTTCAACCAACAGGGTCCCTTAACAGGAAAGCTAGCTTGTATTCACCCGT	1772
QY	1754	TGGCGGAACAAATCTGCTGTAATGACTCAACAAAGTAGAGGCAAGGAGATCAAAATCGAGC	1813
Db	1773	TAGCTGACCAAGTGTCTGTGATGACTCAAAAGGGCGCGAGAGCGATACAAAGGTAGAGC	1832
QY	1814	CATACACGATTAAGGTCAATTGTATCAAGAAAGGAGCGGCGGTCCCTGTTCAAGACTTCAGG	1873
Db	1833	CATACCAAGATAGATTTATCTGTCCCACTGGTACAGCTATACAAATCCCGATTTCCAGG	1892
QY	1874	CATTGAGTAGAGCGCTACGATCGTTTCAACGAGAGGAGTTCTGTAAACGATATCCTGC	1933
Db	1893	CTCTGAGTGAAGTGCAACCAATGATTTTAACGAAACGGGAGTTGTTAAACGTTACTTAC	1952
QY	1934	ACCAATATGGCAATCAACGAGAGACCGCTAAACACTGACGAAGAGATCATATTAAGCTGTAA	1993
Db	1953	ACCAATATGGCGTTAAACGAGAGGCGCATTTGAATTAACGATGAAGAAAGATCTCAAGGTTGGA	2012
QY	1994	AGACTCAAGACACAGACTCAAGATTAACGCTTCGTGATTTGACGCAAGAAAGTGTTTAAGC	2053
Db	2013	AAAGCACTGAGACAGACTCTGATAGTACGTATTTGACATGACGCAAGAAAGTGTGTAAAG	2072
QY	2054	GAGAAAGCGCAGGTCCCTTGTGCTAACCGGTATCTGTGATTCAGTACCAATTTACGAGT	2113
Db	2073	AAGGGGAATGCCGGAACCAATGTGCTGTGCGCGCAACTGATTTGACCCGCAATTCACGAAAT	2132
QY	2114	TTGCGTACGAGAGTCTCAAGACAGACACGAGACACTCAAAAGTCCCAACCATTCGGAG	2173
Db	2133	TTGCGTACGAGAGTTTAAAAACAGTCTCTGCTGACACACAAAGTGTCTACATCCGAG	2192
QY	2174	TCTATGAGTGCACAGTTCAGGTAAATCTGGAATCATCAAAAGCGCTGTGCTTAAGAAAG	2233
Db	2193	TCTATGAGAGTCCAGAGTCTTGAAAGTGTGTATTAATCAAAAGCGCTGTATCCAAAGCGTG	2252
QY	2234	ATCTGCTGTGATGCTCGAAGAAAGAAACCTGCCAGAAATCATAGGAGTGTAAAGAGGA	2293
Db	2253	ATCTGCTGTGATGCTCGAAGAAAGAAATTCATAGGAATCATCAAGACCTCAACGCCA	2312
QY	2294	TGAGACGTATGGATTTGCTGCTAGAGACTGTGCAATTCAGTCTCTTAATGGGGTTAAGC	2353
Db	2313	TGCGGGCATGGAACATGCCGCCCGCAACATGGAATTCAGTGTCTGTAAATGGGGTTAAAC	2372
QY	2354	ACCCGCTTAACACTCTGTACATTGATGAGCAATTTGCTGCCATGCAAGGACGCTGCTGG	2413
Db	2373	ACTCCGTGCAACACTGTACATAGACGAGGCAATTCGCTTGCATGACAGGGACTCTGTCTAG	2432

OY	2414	CACGTGATGGCAATCGTCAAACTTAAGAAAGTGATTTGTGGGGGAGCCCAAAACAATCGC	2417
Db	2433	CACCTATGCCATCGTCAAGCCAAAGAAAGTGATTTGTGGGAGATCCGAAACATTCGC	2436
OY	2474	GCTTCTTTAACATGATGTGCTCGAAGATACATTTTAACATGACATATGACCTGAAGTGT	2533
Db	2493	GCTTCTTTAACTGATGTGTCTTAAGTGTCATTTTAACAAGATATATGACAGAAAGTGT	2555
OY	2534	ACCATAAAGCATCTCTAGAGAGTGCACACAGACTGTAAACGGCATCGTCTCCAGCTCT	2599
Db	2553	ATCAACAAGATATCTCTGGCGATGCATTAAGACAGTGCATCCATTGTCTCCACCTGT	2612
OY	2594	TCTACGACAAAGGATGAAAGACGGTTAACCATGTGCTGATTAATATCATATGATACCA	2653
Db	2613	TCTAAGATTAACGGATGAGAACTGTCAACCCATGCAATGATTAATATGATACCA	2672
OY	2654	CAGGAGCCCAAAAGCCGCAACAAAGATGTGATTTCTAACCTGTTTGAAGATGGGTTA	2713
Db	2673	CCAGTACTACCAACCTTTAAAGATGACATATATTAACCTGTCTCAGAGGGTGGGTTA	2732
OY	2714	AACAGCTACAGATTGACTACAAAATTCAGAAATCATGACTCGGCTGCATCGCAAGAC	2773
Db	2733	AGCAACTGACATGTGACTTAACAAGACAGAGATCATACCGCAGCGGCCCTCACAGGGT	2792
OY	2774	TTACGGCGAAAGCGCTTTATGCTGTACAGTACAAAGTCAACGAGATCACTATCTCGC	2833
Db	2793	TTACTAGAAAGGGGTATACGACAGTGGCCCTCAAGGTCAATGAGAACCCGTTATACGAC	2852
OY	2834	AGACTTCTGAGACGCTGAACGTGTTACTTACACGCAAGAAAACGCTTGTCTGGAAGA	2893
Db	2853	AGACATCTGAGCATGTGATGATGTACTACTTACACGCAAGAAAAGTATGATAGGAAGA	2912
OY	2894	CGTACGTGTGATTCCTCGATTAAGACACTTACAGTAAATATCCCGGAGATTTCACGG	2953
Db	2913	CTTTGGCCGGTGAACCTTGATCAAGAGCTTGACAGCATCGTACCCGGGTAAATTCCACGG	2972
OY	2954	CTTCACTTGGAAGCATCGGCAAGCCGCAACACGACGCATTATGGCACGCTTCTGTAAGC	3013
Db	2973	CCACACTGGAAGAAAGGCAAGCTGACGTGACGTACTATGGCGAAAATTACTTGAGACAC	3032
OY	3014	CGCAGACAGCTGATGTGTTCCAGAAATAGTGAAAGCTGTGCTGGCGCAAGCCTTTAGAGC	3073
Db	3033	CAGCTAGCATGACGTTTTCCAAAACAAAGTGAAAGTGTGCTGGGCCCAAGCCGCTAGAAC	3092
OY	3074	CAGTCTTGGCCACCGCCCAACATTGTGCTGACGAGACACAGTGGGAGACGTTGCAACCAT	3133
Db	3093	CTGTGTGGCCACCGCCCAATATTAACGTGACCCCGCTCCGAGTGGGAGACTTTCAGAGGT	3152
OY	3134	TCAGCATGACAGAGCGTATCTCACTGAATATGGCACGTAACTTTTGTGACCAAGGTTCT	3193
Db	3153	TCAAGGATGACAAAGCGTATTCGCTGAGATGGCCTTAACCTTTTCTGCAACAGATTCCT	3212
OY	3194	TTGAGATGACCTGACAGTGGGTATTTTCCGCTCTACCGGTGCACTTACTTACAGGG	3253
Db	3213	TTGGGGTGCATTCACAGCGGGTGTCTCCGCGCCAACTGTTCGCTGCTTACACCA	3272
OY	3254	ATCAGCACTGGGATTAATCTCGCCAGGGAGAAACATTTATGGGCTTAATATGAGGATAGCAA	3313
Db	3273	ATGAACACTGGGATTAATAGCCAGGTCCAAACATATATATGGGTGTGCAATGGCCACGCTTA	3332
OY	3314	AGGAGTTGTCAACGGGATATCCGTGATCACAAAAGCGGTTGACACAGGACGGGTAGCTG	3373
Db	3333	AAGAACTTGACAGTGGTATCTTGTATTTCTGAAGACGGTGGATACAGATGAGTGGTGT	3392
OY	3374	ATATATAGAAATATATCCATCAAGACCTACTCTCCAACTTAATATGTGGTTCATTAATATC	3433
Db	3393	ACGTTTCGACAGACCTATTAAGACTTAACCCCGCTATATAATGTGTATCCCTCGAATA	3452
OY	3434	GCCGGTGGCCCACTGTTGATCGTGTGACCAAAAGCAGGGTTCAACTGATCACAGCG	3493
Db	3453	GAAAGCTCCCACTCATTTGGTGTTCACACAGAGATACATGGGAACGGTATTAATCTCC	3512
OY	3494	GATTCCTATCTTAAGATGAAGGCAATCTGTGTGGTGTATCGGCACTCTACAGATTC	3553

Db	3513	AGCTAGTACCAAGATGACAGAAAAAACCGTACTAGTAGTGGGATACACTTATGAAACAATAC	35727
OY	3554	CAGGGAAGAAAGTAGAGTCCATGGGTCCATTGGCCACTTAATACCATCAGGTGTGATCTCG	36133
Db	3573	CAGGAAAGAGTGGAGACATTAGCCCAAGCCACAATGACATATTAAGCCGAATCTGG	36522
OY	3614	ATTGGGAATACCTAGCCATGTCCGTAAATAATGACATTAATCTTTGTCAATTTTAGAACCC	36737
Db	3633	ACCTGGGCAATCTGGCCGCTTTTAGGCAAAATGACATCATTTTATTAACGTGAGAACCTC	36922
OY	3674	CGTACAGGAACCATCACTAACCAACAGTCCGGGATTCACGCTATTCACCAACGATGCTTA	37333
Db	3693	CTTACCGCACACCATTTACCAACAGTCCGGGACCAATGCGATCCACCAACGATGCTTA	37522
OY	3724	CGTGAAGGCTGTCCACCACTGAAACACTGACGGGAAACATGTGTGGCTTAAGGATGGCC	37933
Db	3733	CTAAGAAAAGCTGGACCATTTTGAACAAAGCCGGTACGTGCATGCAATGGGCTATGGGA	38122
OY	3794	TTGCTGATCCGGCAACCGAAGATATCATCTGCGGTGGCAACGCTCATTTTAGTTTACC	38533
Db	3813	CTGGGGAACAGACCGAAGAACTTATCTCTGAGTGGCTCGCTCATATTCAGGTTCTAC	38722
OY	3854	GTGTCTGTACCTTAAGAACCTGCGGAAATACTGAGGTTCTCTTGTGTTCTTCGGCA	39133
Db	3873	GTGTGTGCCAACCGAAGTGTGCTTGGGAAACACTGAGGTGCGCTGTGTTTTCGGCA	39322
OY	3914	AGGACAAACGGCAACCAACATACACAGGACACATCCGGTGTAGTGCTTGACAAACATCT	39733
Db	3933	AGGACAAAGGAAACCATCTCCAGATCAAGTATGATGTGAGTGTGTTAACAACATAT	39922
OY	3974	ATCAAGGGTCAACCAAGTACAGGACAGGACGCTCAGCCTACAGATGATCAGAGGTG	40333
Db	3993	ACCAAGGGTCAACTCAATATGATAGCTGGACAGACCTGGCTATAGAGTGGTGGCCGG	40522
OY	4034	ACATTAGCAAGGCGCTGACCAAGCTATCGTTAATGCTCTAATAGCAAGGTCACACAG	40933
Db	4053	ACATTAACAAAGAGCAATGATAGGTTTATTGTTAACGGGCGCAACAACAAAGGCAACCG	41122
OY	4094	GTTCCGGAGTGTGGGGTGCATGTACCGAAATGGCCGGCGCTTTGATGACAGCA	41533
Db	4113	GTGGGGGTGTGTGGCCCTTTACAGAAAGTGGCTTGGAGCTTTGATTAAGACGGCG	41722
OY	4154	TAGCTGTCCGACCGGCTAGACTTGTGAAAGCAGAACCGCTCATCATCATGCTCTAGGAC	42133
Db	4173	TAGCAACTGGTAAGCCGACCTTGTCAAGCATCTCGGAAGCTATCATATGCCGTTGGCC	42322
OY	4214	CCAAATTTTCTAAGATGCCGGAACCGAGGGCGACTTAAGCTCGACGTGCTTACATGA	42733
Db	4233	CTAATTTTCTCGGCTATCAGAAAAACAGAGGACCGAATAATGTCTGAAGTGTATCATGG	42922
OY	4274	GCATAGGCTCATGTGTCAACGCTGAGGGATTAACAAAATATACATACGTAACCGTACTGCA	43333
Db	4293	ACATTGCCAATTAATCAACAAGAGGTTTACTTAAGTCTCCATTCCTGTTTATCTTA	43522
OY	4334	CCGGCATCTATTCTGTGGCCAAAGATGAGTGATGATCATCTGATCACTGTGTCACTG	43933
Db	4353	CCGGCATTTACGACAGGTGTGAAGACAGGGTATGCAATGCTGTAAACATTTTATTCAG	44122
OY	4394	CTTTGCACTACGGAATGCCGATGTCAACATATTTGCTTGGATTAACATGGGAGACA	44533
Db	4413	CCATGGAATTAACCGACGACGACATCATCTATTACTGTCTAGATTAAGCAATGGGAGTCAA	44722
OY	4454	GGATTAATCGAGCCATTCAACGCAAAAGAACGCTGCA---ATTCTGATGATGACAAAC	4510
Db	4473	GAATTAAGGAAGCTATCACTCGAAGGAAGACGTTGAAGAGCTTACTGAGGATTAACGAC	4532
OY	4511	CAGTAGACATTTGATGGTCAAGGCTCACCAACAGCTCTTTGGACGGACAGCAGTT	4570
Db	4533	CAGTTGACATTTGAATGTGTACGGGTACACCCGTTGACAGCTTGGCAGGTATGACTGTGTT	4592
OY	4571	ACTCCGTCAATGAGGGCAAGTTGTATTCATACCTGGAAGGTACAGATTCATCAACCG	4630

Db	4593	ATTCAACCAACCGAGGGCAAGGTGTAATTTCGTACTAGAGGGCACTAAGTTTCACTAACTG	4652
Qy	4631	CCAAGGACATTGTCGGAATTCATGCAATGTGGCCCCCAACAACTGAGGCTAATGACAGA	4690
Db	4653	CCAAAGACGTAGCTGAGATTAAACGTATGTGGCCTTAAACAGAAAGAACCAACGACAGA	4712
Qy	4691	TTTGGCTGTACATCTCTGGGGGAGATATGTCCAGATTCCTGCTCCAAATGCCAGTAGAGG	4750
Db	4713	TTTGGTATTACGTGTGTGGGAGAGATGAAGACAGATCCGCTTAAAGTGTCTGTGAAG	4772
Qy	4751	AGTCAGAGCGCGTCGCTCACCTCAACACATTCACATGCGCTGTAATTAACGTATGACGG	4810
Db	4773	AGTCGGAAGCCTCTTCCCCCTCACACATTCCTCGTGTGTGCACTATGCAATGATCTG	4832
Qy	4811	CTGAGCGCGTATACAGCTTGGCGCTCTCGGAAGAAAGAACAGTTCGCCGTATGCTCATAT	4870
Db	4833	CAGAGCGAGTTTACAGATTAGGTATGGCAAGAAAGAACAAATTCGCAATTTGTTGTCCT	4892
Qy	4871	TCCTGTTCCGAAAGTACAGATTCACAGCGGTGCAGAGCTTACGTGACGACAAACAGTCC	4930
Db	4893	TTCAAGTATCCAAATACAGATTACAGGGGTTCAGAAATTCATGACAGTAAACCTTGA	4952
Qy	4931	TGTTTTCAGGGCGGTGTACACCGGCTGTACACCCAGAGAGTAGGC	4976
Db	4953	TATTCCTCGGACTGTACCCCGGCATATCAATCAAGAAATTCGC	4998
RESULT 6			
VEUS5362			
LOCUS	VEUS5362	11441 bp	RNA linear VRL 12-JUL-1999
DEFINITION	Venezuelan equine encephalitis virus strain 83U434, complete genome.		
ACCESSION	U55362		
VERSION	U55362.2	GI:5442471	
KEYWORDS			
SOURCE			
ORGANISM	Venezuelan equine encephalitis virus		
REFERENCE	Venezuelan equine encephalitis virus		
AUTHORS	Alphavirus; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Weaver,S.C., Salas,R., Rico-Hesse,R., Ludwig,G.V., Oberste,M.S., Boshell,J. and Tesh,R.B. Re-emergence of epidemic Venezuelan equine encephalomyelitis in South America. VEE Study Group Lancet 348 (1992), 436-440 (1996)		
TITLE	1 (bases 8387 to 10063)		
JOURNAL	96332169		
MEDLINE	8709783		
PUBMED	2 (bases 1 to 11441)		
REFERENCE	Wang,E., Barrera,R., Boshell,J., Ferro,C., Freier,J.E., Navarro,J.C., Salas,R., Vaequez,C. and Weaver,S.C. Genetic and phenotypic changes accompanying the emergence of enzootic subtype IC Venezuelan equine encephalitis viruses from an enzootic subtype ID progenitor		
AUTHORS	J. Virol. 73 (5), 4266-4271 (1999)		
TITLE	3 (bases 8387 to 10063)		
JOURNAL	Weaver,S.C. Submitted (18-APR-1996) Pathology, Univ. TX Med. Branch, 301 University Ave., Galveston, TX 77555-0609, USA		
MEDLINE	4 (bases 8387 to 10063)		
PUBMED	Direct Submission		
REFERENCE	Submitted (23-OCT-1997) Pathology, Univ. TX Med. Branch, 301 University Ave., Galveston, TX 77555-0609, USA		
AUTHORS	Weaver,S.C. Submitted (12-JUL-1999) Pathology, Univ. TX Med. Branch, 301 University Ave., Galveston, TX 77555-0609, USA		
TITLE	Sequence update by submitter		

On Jul 12, 1999 this sequence version replaced gi:3849808.

COMMENT	FEATURES
On Jul 12, 1999 this sequence version replaced gi:3849808.	Location/Qualifiers
1. 1141	
/organism="Venezuelan equine encephalitis virus"	
/mol_type="genomic RNA"	
/strain="83U434"	
/db_xref="taxon:11036"	
45. 1523	
/note="contains nonstructural proteins 1, 2, 3 and 4; in-frame termination codon is read through at low frequency"	
/codon_start=1	
/transl_except="(pos:5679..5681,aa:OTHER)	
/product="nonstructural polypeptidein"	
/protein_id="AAD43359.1"	
/db_xref="GI:5442473"	
/translation="MEKVVHVIDIEDSPFLALORFPQFEVKAQVTDNDHNAARAFS HLSKLIETEVDPSPDITLIDIGASPARMVSKKHYHCICPMRCAEDPERLITVYKLLKKK NCEITDIXEIDKKMKELAAVMSDPDLLETMTCLHDESCVBEQVAYVQVYVADGT SLHQANQGVAVWIGFDTPPMFKLAAYSVYSTNWAEDEVLTALNKIGLCSWDVT ERSRGMSSILRRKYLIKPSNNVLVSFVSTIYHERDILRSNHLSPVFLRKNQVTCG ETIYSQGVYAKRIAISPGYKSPGSAALMDEEGFLCCKATDTLNGERSVPEVCTAY PATLCDQMTGILATVDASDAQOKLVNGALNTVNGSTQNTMYNLIYVAQAAPRA RKAKCEQEDDERPGLIGDRVMGSCAFRRHKITSIYKRPDITLIIVNSPSGFL VLPRIQSNLTLEIGLTRIRKMEERKEPPLITIDEVQEAQKCADEKREAREBELRA ALPEPLADVEEPLEADVNDLQEGAGSVETRGILKTVYSAGEDTYSVAUSDPT VLSEKSLCIHPLEAOVYITVTHSGRKRVAPEHGVAVBEHAIIPVQOFALBESA TIYVNERFVNRYLHLIATGALNTDEEYKTVKSEHEGVALYIDRQCVKKELY TGGLGELVDPPEHPEAYESLRIRPAAPVOTIGYVGVGSGSGIISKATYKDL VNAKNCENCAEIIIDVKOMKGLDVNARTVNSVLNCGKHEVETLIYDEAFACHTLR ALAIIRPKAVILCGDPKQCFPMNCLAKHPNHELCTQVPHKISIRCKTSKTSVYS TLFEDKMRITNPETKIEIDTGGSTPEKODDILITCFRCWVQLOLDVGNENMTWA ASQGLTRGVYAVARKVNNPILVAPTSSENVALLTRTEDRIWKTLAGDPMIKTLTAK YSGFATATIEEMQAEHEALIMRHLIERPDPDVQONKANVMAALPVLITAGIDMTT EONMTYVFEFTDRAHSAEIVLNOICVAFPEGLDLSGFSAPVPLSIRNNHMNNSPSP NMVGLNNEVYRQLSRAEYRPLIRPAVACRYVDNMGTGLRNPDIINLVNARLPHALV LHRNEHQSDPSSFVEKTLKRTYLVGEKLSVGEKYDMLSDQEARFARLDLIGPG DVPEYDILFTINVRTPYKTHYVQCEBDATLSMLTKYKACHLNGGTCVSIGVYADR ASEBIIIGAIAROFKFSVCKPKSLBETEVLPVFIGDRKARTHPYKLSLTYNITAT GSRILHEAGCAHVVHVGDIATATGCVIINAANSKQPGGCGALYKKEPESDLOP IEVKARLVKGAAGHIIHAVGPNNTKSEVEGKOLAEVASIAKLIIVNNYSVAIIP LISGIFSGNKDRLTOSLHNLITLADTDDADVALICDCKWEMTLEKVAAREAVEIBI CISDSSVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVDARTTPEPPIIIIEEESBASIILSGPHYOVLVAA DHPEPVSSSMSNIIPIASDFDVDSILITLIGASVYTSAASETNYSYARSMEFLA RPVAPRTVPRNPHAPRPTRTSPDLSASISRTSLVSTSPGPNRYITRELBAREI RAPSRSARTSLVSNPGRVRYITREFEAFVAQOQKRPAGAVITSSDQCGHLOOK SVRTQVTEPDAELVHNRKSLAGRKGYSTDGKTFSTLECTKHPQAOKDAEINAA MMPEATBANQVCWYILGESMSSIRSCPEEESASTPSTLPCLCHANTPEVQQL KASPEQITVCCSFPLEPKRYIRITGVOKQSCQPLFSKRYAVIHPKRYALTOVET PEPTEENQSVEGTEPEPAPVAVD	

GNLGRTRMYRMHGTIEIEIPLHVOVSLHTRPCHIVDGHGYFLARCPAGDSITMEFKK
DSVTHSCVPEYEVKNPVRRELYTHPHBHEAOQOVYAHQONGAIVENHMLPGSEV
DSIVHSJSGSVTVTPPAGTSALVECEGKRISETTITAFQOSCTKEECORARLO
NDKMYNSDKLPKAGATLKGKHLVPLAAGKCTVAPBPMITPGRSISLKHYPK
NPLYTLTROLDPEHYTHELISEPARVPTTEKMEVWENHNPMPKRWAPBTPGNP
HGLHEVITTHYHRYPMSTIGLSICAIIVSVASVATWLNKSVSLITRYLTPNA
RMPICLAVLCCARTARAEITWESLDHNNNOQMFIDLLPLAALIVTLILKVCVC
VPLVAVAGAAGAYEHATTMPESQAGISYNTIIVRAGAYLPLISITPTKILPYN
LEVYCHYKTMDSBPALKCCGSOEPTPYRDEOCKVTGYVPEWGAAYCECDENT
OVSAKYVMKSDCLADHAEYKAHTASVQALNITVGHSHIVTITYVNGETPVANNGV
KLITGPISTAMPDPKRYVOYAGIYNDPFEYAGGQAGADIOSRVSDDIYANT
NLVIOBPKAGAIHPYTOAPSGFBOMKDKMSLEKFTAPFCGLIYNIRANCAVGS
IPLAFDIPDALFTRVSEPTLSAECTINECVSSDFGAIATVKSASKSCKAVHP
SGATLKEAAVELTEOGSVTHIFSTANHPEFRLOICTSYVTCGKDCDCHPMDIVTHP
QYHATFPAASXTAMTWLSTLGGSAVYIIIGLIVATIVAMVYLTNQKNH"
BASE COUNT 3224 a 2825 c 2908 g 2484 t
ORIGIN
Query Match 42.7%; Score 2133.4; DB 14; Length 11441;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3229; Conservative 0; Mismatches 1736; Indels 24; Gaps 3;
QY 10 ACTAATGCATCCAAATATGGAAGAATTTCACCTTGACCTTAGTCGACAGCCCGATATC 69
Db 30 AATTACCTAACCAATGAGAAAGTTACCGTTGACATCGAGAGAACAGCCCATCTCTC 89
QY 70 AAGTCGTACAGCGGAGCTTTCACAACTTTGAGATCGAAGCAAGGCAAGTCACTGACAAT 129
Db 90 AGAGCTTTGCAACGAGAGCTTCCCGAGTTTGAGTAGAAGCCMACAGATCGATATAT 149
QY 130 GACCATGCAATGCGAGAGCGTTTTCGATGTGCGAACAAAGCTCATTTAGAGCGCAATC 189
Db 150 GACCATGCTAATGCGAGAGCGTTTTCGATGTGCGAACAAAGCTCATTTAGAGCGCAATC 209
QY 190 GACCGGAGCAAGTTATCTTGACACTTGAAGTGGCGCCCTGACAGATGCATTTCCAT 249
Db 210 GACCATGCGACAGATCTTGACACTTGAAGTGGCGCCCTGACAGATGCATTTCCAT 269
QY 250 CACCGCTATCATTTGATCTGCCCTGATGATTAAGCGCTGAAGACCGGACAGACTCAACAG 309
Db 270 CATAGATTCATTTGATCTGTGCGATGATGATGATGATGATGATGATGATGATGATGAT 329
QY 310 TATGCAAGAAAGCTTAAAGAAAGT-----GACATTTACGACAAAGATAGCCCTTAAG 363
Db 330 TATGCAACTAAGCTGAAGAAAGTAAAGTGAAGAAATTAATGACAGAGATTTGACAAAGAA 389
QY 364 GCGGCAAGCTGCTGGAAGTATGTCACACCAAGCAAGAGATCTCATCTTGTGTATG 423
Db 390 ATGAAGGAGCTGCGCGCGCTGATGAGCGACCTTGACCTGGAAGCTGATGATGCTC 449
QY 424 CACACAGAGCGCAGGTGAGTACTTTGGAAGTGAAGTGAAGTATCCAGATGTGATGCGCA 483
Db 450 CACACAGAGCGTGTGCTGCTACGAGGGGCAAGTGTGCTTTACAGAGATGATACCGCA 509
QY 484 GTTCATGACCGGACATCATCTACACCAAGCGCTTAAAGAGTTAGACAAATTTACTG 543
Db 510 GTTGACGAGCGGACAGTCTGATACCAAGCAAGAGAGTTAGAGTGTGCTGCTACTG 569
QY 544 ATAGGCTTTGACAGACCCCTTTTATGTAAGAAAAATGAGAGTTCTTACCTTACTTAC 603
Db 570 ATAGGCTTTGACAGACCCCTTTTATGTTAAAGAACTTGCTGAGCATATCTCATATAC 629
QY 604 AACAGCAATGGGCTGACGAGAGATTTGGAAGCAGCTAATGGCTGCGATACCTCA 663
Db 630 TCTACCACTGGGCGGAGAAACGCTGTTAAAGGCTGTTAATAGGCTTATGAGCTCC 689
QY 664 GATCTTACAGAGAGCGCTTGAAGAACTCTCAATCTTTAGAAAGAGGCTCCACCT 723
Db 690 GACGCTAGGAGCGGTACGTAGAGGATGTCATTTTGAAGAGAAATTTTGAACCA 749
QY 724 ACTAATTAAGATCATTTTGGGTGTTCAACATTTACACAGAGATGATCACTGTTA 783
Db 750 TCCAATATATGTCATTTCTGTGTGCTGACCATCTTACACAGAGAGGAGATTTACTG 809

QY 784 CGTAGCTGCACTTCTCCAAACGTGTTCCAATTGAAGAAAGTCTAATCTTCACAGGTAGA 843
Db 810 AGGAGCTGGACACCTGCCCTCTGTATTTTACCTTAGGTGCAAGCAAAATTTACATGTGG 869
QY 844 TGTGGACCAATTTGCTAGCTGGAAGGTACGTCA TCAAAAAGATTAACGATCAGCCAGGA 903
Db 870 TGTGAGACTATAGTTAGTTGCGAGCGGATAGCTCGTTAAAGAAATAGATATCAGTCCAGGC 929
QY 904 CTATACGCTAAGTTAGAAACTTGGCGCTCCCAATGCTATGCCGAGGTTTCTTGAGTTGC 963
Db 930 CTGATGGGAAGCCTTACAGGTATGCTGCCACGATGACCCGAGGAGTTCTTGCTGC 989
QY 964 AAAGTCACAGATFACGCTGCGCGGAGAGGATTTCTTTGTGTGTGATACGATATGACA 1023
Db 990 AAAGTCACAGACATTTGAACGGGAGAGGATCTTTTCCGCTGTGACGATATGTGCA 1049
QY 1024 GCCACACTTTGCGATCAGATGACAGGATTTGSCAACTGACGTTAGTGTGATGACGCA 1083
Db 1050 GCTACATTTGTGTGACCAATGATGCTGCACTAGTCCAAACAGATGTGAGTGGAGCGAGCG 1109
QY 1084 CAAAACCTATTGTTGGGCTTAACCAAGAAATTTGCTCATGTGTAGAGACCAAGAAAT 1143
Db 1110 CAAAACCTGCTGTTGGGCTTCAACAGGAGCATGTGTCACAGGTGCAACCAAGAAAC 1169
QY 1144 ACTAACCAATGCGAGAACTATCTATTAACAGTGTGCGCCAGGCGTTTCCAGGTGGGCG 1203
Db 1170 ACCAATATACATGAAGAAATTTATTTTGTGCGCTGATGCGCCAGGATTTGTAAGTGGCA 1229
QY 1204 CGTGAACATCGTGGCGACTTGAGCAGAGAGAAAGATAGGAGTGGCGGAGCGCACTCTT 1263
Db 1230 AAAGAAATTAAGAAAGATTAAGAAAGATGAGAGGCGCACTAGACATACGATAGACAGTTA 1289
QY 1264 ACTATGGCTGCTGCTGGGCTTTCAAGACCAAGAAATCACATCATCTACAGAGGCT 1323
Db 1290 GTTATGGGCTGCTGCTGGGCTTTTGAAGGACAAAGATTAATCTATTTATTAACGCGCCA 1349
QY 1324 GGTACGCAAAACAAATTAAAGAAAGTACCTGCGCTTTGCTCATTTGATTTCCAGGCTT 1383
Db 1350 GATACCCAAACATCATCAAAAGTAAACAGGATTTCCATTTGCTGCTCCAGGATTA 1409
QY 1384 ACCAGCAAGGCTCGATATGAGGCTTCCGCGTAGGCTCAAGCTGCTGTTGAACCACT 1443
Db 1410 GGCAGTAAACATGATGAGATGAGGCTGAGAACAAAGATTAAGGAAATGCTGAAGAGCGC 1469
QY 1444 GTCAAAACCGGACCGGCTATTACATGAGCGGATGTGAGCATCTGCGTGTTCACAGCA 1503
Db 1470 AAGGAGCGGCACTCTCATTTACTGACGAGGACGTACAGAAAGTTAAGTGGCAGCGAT 1529
QY 1504 GAGCTGAAGAGTGGCTGACAGCGGAGAGATCAAGAGAGCCCTGCAACCTTGTCCCT 1563
Db 1530 GAGGCTAAGAGTGGCTGAGAGCGGAGAGTGGCGCTGACATCAACCTTTTGGCAGCT 1589
QY 1564 GAAATTAAGAAAGAGACCGTGAAGAGGAGAGTAACTCATTTATGCAAGAGGAGAGCA 1623
Db 1590 GATGTTAGAGAGCCCACTGAGAGCGGATGTGATGATGATGATGATGATGATGATGATGAT 1649
QY 1624 GGTAGCTGAGAGACACACAGAGACATCAGGTTGACAAATTACCAAGCGGAGAGAG 1683
Db 1650 GGCTCAGTGGAGACACCTCGGTGATTAAGGTTACCAAGCTATGCGCGGAGAGAGCAAG 1709
QY 1684 ATTGGGCTTAAGCTATACCTTACCCAGGCGGTTATGGAATAGTAAAAATCTGCGGT 1743
Db 1710 ATCGGCTTTTACGAGTGTCTTTCTCCAGAGGCTGTACTCAAGATGAGAAATCTATCTTGC 1769
QY 1744 ATCCACCATTTGGCGGAGAAAGTACTGTATATGCTCAAAAGGTAGGCGAGGAGATAC 1803
Db 1770 ATTACCTCTCGGTGAAGAAAGTATAGTATTAACACATCTGCGGAGAAAGCGGCTTAT 1829
QY 1804 AAAGTCAGGCAATACACGCTAAGCTATGTATACAGAGGAGAGCGGCTCTGTTCAA 1863
Db 1830 GCCGTGAACCTTACATGAGAAAGTGTGTGTCAGAGGAGACATGCAATACCGTCCAG 1889

QY 1864 GACTTCAGGATCTAGTGAAGCGCTACGTCGTTTCAACGAGGAGGATTCGTAAAC 1923
Db 1890 GACTTTCAAGCTCGAGTGAAGCGCCACCATTTGTATCAACGAGCGAGATTTGTAAAC 1949
QY 1924 AGATACCTGCACCAATCGCAATCAACGAGAGCGCTAAACACTGACGAAGACTACTAT 1983
Db 1950 AGGATCTGCGACCAATTTGCGACACATGAGAGCGCTGAACACAGATGAAGAAATTTAC 2009
QY 1984 AAGCTGTAAAGATCAAGGACACAGACTCAGATATACGCTTCGATATTGACGACGAAG 2043
Db 2010 AAAAAGTCAAGCCAGGAGACACGACGCGCAATCTGTATGACATCGACAGAAACAG 2069
QY 2044 TGTGTAAAGCAGAAAGACGAGGTCCCTTGTGCTTAACCGGTGATCTGTAGATCCACA 2103
Db 2070 TGGCTCAAGAAAGAACTAGTCAAGGGCTAGGGCTTAAGCGGAGCTGTGTGATCTCTCC 2129
QY 2104 TTTCAAGATTTGGCTGACAGAGTCTCAAGACGACGACGACCTCACAAAGTCCCA 2163
Db 2130 TTCCATGAAATTCGCTTACGAGAGTCTGAGAAATACGACGCGCTCTTACCAAGTACCA 2189
QY 2164 ACCATCGGAGCTATNGAGTGCAGGTTCAGTAAATCTGGAATCATCAAAAGGCGTGTG 2223
Db 2190 ACCATAGGGGTATNGCGTCCGGGATCAGGCAAGTCTGCAATCATTAAGAGCGAGTGC 2249
QY 2224 ACTAAGAAAGATCTGTTGTGAGTGCAGAAAGAAAGAACTCGGCAAAATCATCAGGGAT 2283
Db 2250 ACCAAAAAGATCTGTGTGTAGGGCCCAAGAAAGAACTCGGCAAAATTAAGAGAC 2309
QY 2284 GTAAGAGAGTGAAGCTATGATGTGTCTGTAGAGCTGTGCAATTCAGTCTCTTAAT 2343
Db 2310 GTCAAGAAATGAAGAGGCTGAGCTCAATGCAGAACTGTGAGCTCAGTCTCTTGAAT 2369
QY 2344 GGGGTAAAGCAACCCGTTAAACACTCTGTATGATGAGGCAATTCGCTGCAAGGAGG 2403
Db 2370 GGATGCAAAACCCCGTATGAGACCTGTATATGACGAGCTTTGTGCTGTGACGAGGC 2429
QY 2404 ACGCTGTGCGCATGTATGTCATCGTCAAACTTAAGAAATGTATGTGCGGGAGCCCA 2463
Db 2430 ACTCTAGAGCGCTCATAGCATATAGACTTAAGAAAGAGTGTCTGCGGGAGCCCA 2489
QY 2464 AAAACAATGCGGCTTTTAAATGATGTCTGTAAGTACATTTTAACTGACATATATGC 2523
Db 2490 AAAACAGTGGGCTTTTCAATATGATGTGCTGAAGTGAATTTTAACTGAGATTTGC 2549
QY 2524 ACTGAAGTACCAATTAAGATCTATGAGAGTGCACACAGCTGTAACGCCATGCTC 2583
Db 2550 ACGGAGGCTTTCACAAAGATCTCTGCTGTGCACTTAATCTGTGATCTTGGTGTGC 2609
QY 2584 TCACGCTCTTCTACGACAGACGATGAAGACGCTTAACCATGTGTGATTAATATATC 2643
Db 2610 TCAACCTTTGTTACGACAAAGATGAAGACGACGAAACCGAAAGACATTAAGTTGAG 2669
QY 2644 ATAGATACACAGGACACAAAGCGCACAAAGATGATCTGATTTTAACTGTTTCA 2703
Db 2670 ATTACACTACTGCGAGTATTAACAAAGACGAGCTCATCTCATCTTGTTCGA 2729
QY 2704 GGATGGGTAAGACGCTACAGATTTGACTCAAAATTCGAAATCATGACTGCGGCTCA 2763
Db 2730 GGGTGGGTGAAGAGTTGCAATATGATTAAGGCAACGAAATATATACGCGAGCTGCC 2789
QY 2764 TCGCAAGACTTACGCGGAAAGCGTTTATGCTGTCAAGTACAAAGTCAAGAGATCCA 2823
Db 2790 TCTCAAGGCGTGAACCGTAAAGGGGTGTATGCGTTGGTACAAAGTGAATGAATGCC 2849
QY 2824 CTCTACTGCAAGCTTTGAGAGCTGAAGCTGTTTCTTACACGACGAAAAACGATTT 2883
Db 2850 CTGTACGCAACCACTCAGAAACATGTGAACGTCTACTGACCCGACGAGAGACGTATC 2909
QY 2884 GTCTGGAAGAGCTAGCTGTGATCTCCGATTAAGACATTAACGCTTAATATCCGGG 2943
Db 2910 GTGTGAAAAACATGACCGGCGATCTCATGTGAATAAAAACATGACGCGCAAGTATCAAGG 2969
QY 2944 GATTTCAGGCTTCAATTGAGCACTGTGACGCGCAACGACGCCATTATGGACGCGTT 3003

Db 2970 AATTCACTGCGACGATAGAGAAATGGCAAGCAGAGCATGAGGCGCATATGAGCACATC 3029
QY 3004 CTGTATAGCCCGCAGACAGCTGATGTGTTCCAAAATTAAGGTGAACCTCTCGCGCAAG 3063
Db 3030 TTGGAGGACCCGACCCCTTACCGAGTATTCAAAATTAAGCGAAGCTGTGTGGCCAAAG 3089
QY 3064 GCTTTAAGCCAGCTTGTGGCCACGCGCAACATTTGTGTGACGACAGCAGTGGGAGACG 3123
Db 3090 GCTTTGTGCGCGTACTGAGAAATCTGCAGGCAATGACATGACATGAAATGGAACACT 3149
QY 3124 TTGCACCATTTCAAGATGACAGAGCTACTCACTGAAATGGCACTGAACCTTTTTC 3183
Db 3150 GTGATTTATTTGCAAAAGCAAAAGCTCAGCTCAGCAAGATGATTAAGAACCAACTATGC 3209
QY 3184 ACCAGTCTCTTGGAGTGAAGACCTGAGAGTGGTATTTCCGCTCAACCGTGCCTT 3243
Db 3210 GTGAGTTCTTTTGGACTGACCTGAGCTCGGTCTATTTTCTGACACCCACTGTCCGTTA 3289
QY 3244 ACTTACAGGATCAGCACTGGGATTAATCTCGCCAGGGAAGAACTGTATGGCTTAATGA 3303
Db 3270 TCCATTTAGAAATATCACTGGATTAATCCCGCTGCTTAACATGACGGCTGAATTA 3329
QY 3304 GAGGTACAAAGAGTTGTACGCGCATATCTGTGATCAAAAGCGTTGACACAGGC 3363
Db 3330 GAAGTGTCTGCTAGCTCTCCCGAGGTAACCAAGCTGCTGAGAGTTGCTACTGGA 3389
QY 3364 AGGTACTGATATTAAGAAATTAATACATCAAGACATCTCTCAAAATTAATGTGTT 3423
Db 3390 AGAGTCTATGACATGAACATCTGACACTCGCAATTAATGATCCGCAATTAATCTAGTA 3449
QY 3424 CCATTAATGCGCGGTTGCCCACTCGTTATCGTTAGCCAAAGGACAGGTACAACT 3483
Db 3450 CCAATGAACAGAAAGATCGCTCATAGCTTTAGTCTTCAACGTAATGAACCCACAGAGT 3509
QY 3484 GATCAGCGGATCTTATCTATGATGAAGAGGCAATCTGTGTGTGATGCGGATCT 3543
Db 3510 GACTTCTCTTCAATTCGACAAACTGAAGGCGCAACTGTGTGTGCGGGAGAAG 3569
QY 3544 ATCAGCATTCGAGGAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 3603
Db 3570 TTGTCCCTCCAGGCAAAAGGTTGATGTGTGTCAGACAGCTTGAAGCTTCACTTCA 3629
QY 3604 TGTATCTGATTTGGGAATACCTAGCATGTGCGTAATATGACATTAATCTTGTCAAT 3663
Db 3630 GCTGCGTGAATTTAGTATTCAGTGAAGCTGTGCCAAATTAACGACATTTATTAAT 3689
QY 3664 GTTAGAACCCCGTACAGAAACCATCACTACCAAGTGCAGAGATCAGCTATCACAC 3723
Db 3690 GTGAGACTCCATATTAATATCAATCATTAATCAGCAGTGAAGACACGCCATTAACCTT 3749
QY 3724 AGCATGTAAAGTGTAAAGCTGTGTCCACACTGAAACATGGCGGAACATGTGTGCTATA 3783
Db 3750 AGTATGTGAACCAAGAAAGCTGTGTGATCTGATGATCCCGGGAACCTGTGTGACATA 3809
QY 3784 GGGTATGGGCTTGTGATTCGCGCAACCGAATATCACTGCGGTGCGCACGCTCATTT 3843
Db 3810 GGTATGTATAGCTGTGACAGGCGCAGGAAAGCATTTATGTGTGTAATAGCGCGCAGTTC 3869
QY 3844 AGTTTACCCGTGTCTGTACGCTTAAGAACCTGCGCAAAATTAAGTGTCTTGTGTG 3903
Db 3870 AAGTTTCCCGGGATGACAAACCGAAATCTCACTTGAAGAGACGAAGTGTCTTGTGTG 3929
QY 3904 TTTCTTGGCAAGACCAACGCGCAACCAACATGACAGACAGACTCGGTGTATGCTT 3963
Db 3930 TTCAATTTGGTACATGCGCAAGGCCCTGTACGCAATCTTAAAGCTTTATCTTCACTTG 3989
QY 3964 GACAACTATCATCAAGGTCAACCAAGGTACGAGGACGAGAGCTCCAGGCTACAGATG 4023
Db 3990 ACCAACATCTATACAGTTCCAGACTCCAGAGCCGAGATGCCGACCTCATATCATGTG 4049
QY 4024 ATCAGAGTGAATTTAGCAAGCGCTGACCAAGCTATCTGTAATGCTGTAAATAGCAAA 4083

```

Db      4050 GTGCGAGGGGACATTGCCACCGCCACCAAGAGTATCATTAATGCTGCTAACAGCAAA 4109
Oy      4084 GATCAACACGAGTTCCGAGTGTGCGGTGACTGTATCCGAAATGCGCGGCTCTTTGAT 4143
Db      4110 GGACAACTTGGCGAGGGGTGTGCGAGCGCTGTATTAAGAAATTCGCGAAAGCTTCGAT 4169
Oy      4144 AGACACCAATAGCTGTGCGGACCGGCTAGACTTGTGAAGCAAGAAC--GCTCATCATTA 4200
Db      4170 TTACAGCGGATGAGAGTGGGAAAGCGGACGTGTCAAGAGTGACTTAACATATCATT 4229
Oy      4201 CATGTGTAGAGACCAATTTTCTTAAGATCCGGAACCGGAGCGGACCTTAAGTCCGA 4260
Db      4230 CATGCCGTAAGGGCCCACTTCAACAAAGTTTCGAAGTTGAAGTGACAAACAGTTGGCA 4289
Oy      4261 GCTGCCATCATGAGATGCGTCATGTCAGACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db      4290 GAGGCTTATGATCATGCTTAATTAATGTCAGATTAACATTAACAGTCACTAGCCGATT 4349
Oy      4321 CCGCTACTGTCAACCGGACTTATCTGTGTGCAAAAGATCGAGTATGATCATATTCGAT 4380
Db      4350 CCACTGTGTCCACCGGATCTTTCCGGAACAAAGACCGACTAACCCATCATTTGAAC 4409
Oy      4381 CACCTGTCTACTGCTTTTGACACTAGCGGATGCGGATGTCACCATATATTCGTGATTA 4440
Db      4410 CATTGCTGACAGCTTTAGACACCACTGATGAGATGATGACCATATATGACAGGACAG 4469
Oy      4441 CAATGGAGACACGAGTATGAGGACATTCACCGCAAGAAAGGTGGAATTCGTGA- 4499
Db      4470 AATATGGAAATGACTCTCTCAAGAGACAGTGTAGAGAAAGAGTGAAGAGATATGC 4529
Oy      4500 -----TGATGACAAAGCCAGTAGACATTGACTTGTGAGGCTCCACCCAAAC 4545
Db      4530 ATATCCGACGATTCCTCGGTGACAGACCGGATGAGAGCTGTGAGGTATATCCGAAG 4589
Oy      4546 AGCTTTTGGGAGGACAGACAGGTTACTCCGTCAATGAGGAGCAAGTTGATCATACTG 4605
Db      4590 AGCTTTTGGGAGGAGGAGGCTACAGACAGATGAGCAAGATTCATATTTG 4649
Oy      4606 GAAGGTACGATTCATCAGACCGCCCAAGGACATTCGCGAAATCCATGCAATGCGCC 4665
Db      4650 GAAGGACCAAAATTTACCGAGCGGCCCAAGGATTAAGCAAAATTAATGCTGCGCC 4709
Oy      4666 AACAAATCGAGGCTAATAGAGAGATTTGCTTGTATCATCTCGGGAGAGATGTCAGC 4725
Db      4710 GTTGAACGAGGAGCAATGAGAGATATGCAATATCTTGTGAAAGCATGAGCAGC 4769
Oy      4726 ATCCGCTCCAAATGCGCAGTAGAGAGTCAAGGCGTGTCTCACTTCAACACTTCA 4785
Db      4770 ATTAGGTCAAAATGCGCGTCAAGAGTGGAAAGCTTCCACACACTAGACGCTGCT 4829
Oy      4786 TGCCGTGTATTAAGCTATAGCGGTAGAGCGCGCTATACAGGTTGCGCTTCCGAAGAA 4845
Db      4830 TGCTGTGATCATCAAGTATAGCTCAGAAAGATCAACCGCTAAAGCCCTACAGTCA 4889
Oy      4846 GAACAGTTCGCGGTATGCTATCTCTGTTGCGGAAGTACAGAGTACAGAGCGTGCAG 4905
Db      4890 GAACAAATTAAGTGTGCTATCTCTTTCATTGCGGAAGTATAGATCACTGTGTGAG 4949
Oy      4906 AAGCTACAGTACAGCAACCACTCTGTTTTCAGCGGTGTAACACCGGCTGTACACCC 4965
Db      4950 AAGATCCAGTGTCTCCAGCTTACTGTTCTACCTAAGTGTCTGCGTATTCATCA 5009
Oy      4966 AGGAAGTAC 4974
Db      5010 CGGAAGTAC 5018

```

```

RESULT 7
AF004472 11420 bp RNA linear VRL 12-JUL-1999
LOCUS AF004472 Venezuelan equine encephalitis virus strain 66457, complete genome.
DEFINITION AF004472
ACCESSION AF004472.2 GI:5442461
VERSION

```

```

KEYWORDS
SOURCE
ORGANISM
Venezuelan equine encephalitis virus
Virus; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; VEEV complex.
REFERENCE
AUTHORS
Powers,A.M., Oberste,M.S., Brault,A.C., Rico-Hesse,R.,
Schmura,S.M., Smith,J.F., Kang,W., Sweeney,W.P. and Weaver,S.C.
Repeated emergence of epidemic/epizootic Venezuelan equine
encephalitis from a single genotype of enzootic subtype ID virus
J. Virol. 71 (9), 6697-6705 (1997)
JOURNAL
MEDLINE
97404681
PUBMED
9261393
REFERENCE
AUTHORS
Wang,E., Barrera,R., Boshell,J., Ferro,C., Freier,J.E.,
Navarro,J.C., Salas,R., Vasquez,C. and Weaver,S.C.
Genetic and phenotypic changes accompanying the emergence of
epizootic subtype IC Venezuelan equine encephalitis viruses from an
enzootic subtype ID progenitor
J. Virol. 73 (5), 4266-4271 (1999)
JOURNAL
MEDLINE
99214366
PUBMED
10196323
REFERENCE
AUTHORS
Powers,A.M., Oberste,M.S., Brault,A.C., Rico-Hesse,R.,
Schmura,S.M., Smith,J.F., Kang,W., Sweeney,W.P. and Weaver,S.C.
Direct Submission
Submitted (16-MAY-1997) Pathology, Univ. Texas Med. Branch,
Galveston, TX 77555-0609, USA
4 (bases 1 to 11420)
Wang,E. and Weaver,S.C.
Direct Submission
Submitted (12-JUL-1999) Pathology, Univ. Texas Med. Branch,
Galveston, TX 77555-0609, USA
Sequence update by submitter
On Jul 12, 1999 this sequence version replaced gi:3849802.
REMARK
COMMENT
FEATURES
source
1..11420
location/Qualifiers
/organism="Venezuelan equine encephalitis virus"
/mol_type="genomic RNA"
/strain="66457"
/db_xref="taxon:11036"
45..7502
/note="contains nonstructural proteins 1, 2, 3 and 4;
in-frame termination codon is read through at low
frequency"
/codon_start=1
/ciransl_except=(pos:5658..5660,aa:OTHER)
/product="nonstructural glycoprotein"
/protein_id="AAD43356.1"
/db_xref="gi:5442461"

```

CDS

```

/ciransl_except=(pos:5658..5660,aa:OTHER)
/product="nonstructural glycoprotein"
/protein_id="AAD43356.1"
/db_xref="gi:5442461"
/translation="MEKVVHVIDIEDSPFLALQSPFQPEYEAQVTDNDHANARAF
HLAKLLETEDVSDTDLIDISAPARRRYSHKHCICPMCAEDPDLVYATLTK
NFKETIDDELKXKKELAAVSDPDLLETETCLHDDSCREGQAVQVADYAVGPT
SLVHDAKGVAVAYWIGFTTPEFMKLAGAVPSGTMADETVLTAANIIGLSSDM
ERSRGMILKKTLLKSNVNVLSVGSITIEKDLBSMLPSVFLRGKONYCRC
ETIYSCGVYKRIALISGLYKSGSYAATMHRGCTCCXTDTLNGSRVSPVCTY
PATCDQMTGILATDVSHDOKLVLGNIGIVNQRORTNTKATLIPVADAPA
RWAKVEYDEQDERPLGLRDLQVGCWAPRRKHTISYRPDTQITIKVNSDHSF
VLPIRISNTLEIGLTRIRKMLBERKESPLTAEEDVOCACDAEABEABELRA
ALPLADVEBPTLEADVDTLMOAGAGSVETPGLIVTSVAGEDKIGSVASLPOA
VLKSEKSCILPBLEADVDTLMOAGAGSVETPGLIVTSVAGEDKIGSVASLPOA
TIVNBERFVNRVYLAHTAGCATNPDEYKIVPSRDEYLYIDRCKVKEVLA
TGGLGTSELVDPFPHRYEBSLTPRPAAPVPTTGYGVSGSGSITIKASVTKDL
VVSARKENCAEIRIDVKKMKLDVNAKTVDSVLNGCKPEVTLIDAEFCAVGTSLR
ALIAIIRPKKAVLGGDPKQCGFPMWCKLAFHNEICVQVHKQISRCTKSVSVAS
TLFYDKMRITNPKETKLEIDTGTSTKPEODLLTCFRGWKLOLDYKGNEMTAA
ASOGITRGVAVAYKVNENPLVAPTSBVNVLLRTEDRIVMTKLADDPWIKTLTA
YVGNFTATIEEMOAEHAIIMBHIIERDPDTAPONKANVCAKALVPLTKAGIDMT
EQRNTVDYFETDKAHSAEIVLNUCVRFGIDLSGLPSATPVLSIRNNHWDSPP
NMVGLANKEVNSQSRIRPLRAVATGKVVDNMTGTLICNDIPRLVFNVRNLPLALV
LHNHEHPOSDFSSFVSKLGRVTVLVEKSCITVDMVTDWLSDEAPETRLDGIIG
AVPRYDIFVIVNRPYKYHNYOQCEBHAKLSMLTKACHLNFGCTCISYGYADR
ASESITIGAIARQFESRVCKSKSLSETEVLFVIGDRKARTNPNYLSLTINITYT

```


Db 1350 GATACCCAAACATCATCAAGTGAACAGCGATTTCCACTATTGTGCTGCCAGATA 1409
Qy 1384 ACCAGCCA CGGAGCTCGATATGGCTTCCGCGTAGGCTCAAGCTGCTTGAACCACT 1443
Db 1410 GGCACATACACATTTGAGATTCGGGCTGAGAACAAATACGAAAAATGCTAGAAAGCCG 1469
Qy 1444 GTCAAAACCGCAGCGGCTATTCAATGCGCGATGTGGAGCATCTGCGTGGCTTACAGCA 1503
Db 1470 AAGAGCCGTCACCTCTCATTTACTGCCGAGACGTACAAAGATTAAGTGGCCAGCCGAT 1529
Qy 1504 GAAGCTGAAGATGGCTGCCAGCGAAGATCAAGAAAGCCCTGCCACCTTCTCCTT 1563
Db 1530 GAGGCTAAGAGAGTGCCTGAAGCCGAGAGTGGCGCTGACACTACACTTTTGGCAGCT 1589
Qy 1564 GAATATAGAAAAAGACCGTAGAGCGAGAGTAGACCTCATTTATGCAAGAGCGAGAGA 1623
Db 1590 GATGTTGAGAGACCCACTCTGAGAACCGATGTTGACTTGAATCTACAGAGGCTGGGCT 1649
Qy 1624 GGTACCGTGAAGACACAGAGAGACATCAGGGTGAACAATTACCCAGCGAGAGAG 1683
Db 1650 GCGTCAGTGAAGACCTCTGCTGTTGAATAAGTTAACAGCTATGCGCGGAGAGACAAG 1709
Qy 1684 ATTGGCTCTTACGCTATCTTCACTCCAGCGCGTATGTAATATGTAATAATGCGCTGT 1743
Db 1710 ATCGGCTTTAGCAGTGTCTTCTCCAGCGCTGATCTCAAGAGTGAATACTATCTTGC 1769
Qy 1744 ATCCACCATTTGGCGGAAACAACTAGCTGTAATGATCTCAAAAGGTAAGGCGAGATAC 1803
Db 1770 ATTCACTCTCTGCTGAACAACTCATAGATTAACAACCTGCGCGAAAGGCGTAT 1829
Qy 1804 AAAGTCAGCCATACACGCTAAGGTCAATTTGACAGAAAGGAGCGGCTCCCTTTCA 1863
Db 1830 GCCGTGAACCTTACATGAGAAAGATGCTGCCAGAAAGACATGCAATACCCCTCAG 1889
Qy 1864 GACTTCCAGGCTTGAAGTGAAGCGCTACGATCGTTTTCACGAGAGGAGTTCGTAAAC 1923
Db 1890 GACTTTCAGGCTTGAAGTGAAGGTCCACATTTGTGTCAACGAACGAGAGTTCGTAAAC 1949
Qy 1924 AGATACCTGACCAATCGCAATCAACGAGAGGCGTAAACATGAGAAAGTACTAT 1983
Db 1950 AGGTATCTGACCAATATGCTCCACATGAGAGAGCGCTGAACACAGATGAATAATATAC 2009
Qy 1984 AAAGCTGTAAAGCTCAGAGACACAGATCTCAATAATCTTCTGATATTTGACGACAGAAAG 2043
Db 2010 AAAATTGTCAAGCCAGAGAGACGACGCGCAATCTGTATGACATGACAGAAACAG 2069
Qy 2044 TGTGTTAAGCGAAGAGCGAGGTCCCTTTGTGCTTAAACGGTGTGTTAGATCCACA 2103
Db 2070 TGCGTCAAGAAAGAACTAGTCACTGGGCTAGGGCTTACAGCGAGCTGTGATCTCCC 2129
Qy 2104 TTTCAGAGTTTGGCTAGAGGTCTCAAGACAGACAGACGACCTCACAAATGCCA 2163
Db 2130 TTCCATGAAATTCCTACGAGAGTCTGAGAACACACGCGCGCTTACCAATAACA 2189
Qy 2164 ACCATCGAGTCTATGAGTGCAGAGTTCAGGTAATCTGAATATCAAAAAGCGCTGTG 2223
Db 2190 ACCATAGGGGTATAGCGGTGCGGGTTCAGGCAAGTCTGCAATATTAAGCCGCAATC 2249
Qy 2224 ACTAAGAAAGATCTGTTGTGTGATGTCGAAGAGAAACCTGCGCAAAATCATCAGGAT 2283
Db 2250 ACCAAAAAGATCTGTGTGTGAGCGCCAAAGAAAGAAACCTGCGCAAAATTAAGGAGAC 2309
Qy 2284 GTAAGAGAGATGAGACGTATGATGTTGCTGTAGAGATCTGTGATTCAGTGTCTTAAT 2343
Db 2310 GTCAAGAAATGAAGAGCGCTGACGTCAATGCAAACTGTGACTCAGTGTCTTTGAT 2369
Qy 2344 GGGGTTAAGCAACCCGTTAACTCTGTACATTTGAGAGCAATTTGCCATGCAAGG 2403
Db 2370 GGATGCAAAACACCCCGTGAAGACCTGTATATTTGAGAAAGCTTTTGTCTGTATGCAAGC 2429
Qy 2404 ACGCTGTGCACTGATTCGTCCTAAACCTTAAGAAAGTGTATTTGTGCGGGAGCCCA 2463
Db 2430 ACTCTAGAGCGCTCATAGCCATCAATAAGACCTTAAGAAAGCGAGTCTCTGCGGGATCCA 2489

Qy 2464 AAACATGCGGCTTCTTAACATGATGCTGAAAGTACATTTTAACATGACATATGC 2523
Db 2490 AAACATGCGGCTTCTTAATATGATGCTGAAAGTACATTTTAACATGACATATGC 2549
Qy 2524 ACTGAAGTGAACATTAAGACATCTGTAGAGGTGCACACAGACTGAACCGCATCTGC 2583
Db 2550 ACGAGGCTTCTCAACAAAGACATCTCTGCTGTACATTAATCTGTGACTTGGTCTGC 2609
Qy 2584 TCCAGGCTTCTTACAGACAGGAAATGAAGACGTTAACCCATGTGCTATTAATCATC 2643
Db 2610 TCAACCTGTTTAAAGACAAAGAAATGAAGACGAAACCGAAAGACATGAATGAG 2669
Qy 2644 ATAGATACCAAGGAGACCAAAAGCGGACAAAGATGATCTGATTTCACTGTTTCA 2703
Db 2670 ATTGACATCTGCGAGATCTTAACCTGAGAGAGCATCTCATCTTCTGTTTCA 2729
Qy 2704 GGATGGGTGAACAGCTACAGATTGACTCAAAATAACGAAATCATGACTGCGCTGCA 2763
Db 2730 GGGTGGTGAAGCAGTTCAAATAGATTACAAAGGCAAGAAATATAGCGCAGCTGCC 2789
Qy 2764 TCGAAGACATTAAGCGGAAAGCGTTATGCTGTACAGTAAAGTCAAGAAATCCA 2823
Db 2790 TCTCAAGGCTGACCCGTAAAGCGGTATGCTGCTTGGTACMAAGTGAATGAATCCC 2849
Qy 2824 CTCTACGCACTTCTGAGACGTGAACGTTACTTACACGCAAGAAAGCAATTT 2883
Db 2850 CTGTACGACCCACTTGAAACATTTGAAGTCTTACTGACCCGACGAGAGACCGTATC 2909
Qy 2884 GTCTGAAGACGCTAGCTGTGATCTCTGATTAAGACACTTACAGTAAATATCCGG 2943
Db 2910 GTGTGAAGAACATTAAGCGGAGATCATGATTAAGAAACATGACGGCCAGTATCTGG 2969
Qy 2944 GATTTACGCTTCAATGAGACATGCGAGCGGCAACGAGCCATTAATGSCAGCCTT 3003
Db 2970 AATTCACTGCGCAGATGAGAAATGCGAACGAGCATGAGCCATCATGAGCACAATC 3029
Qy 3004 CTGTATAGCGCGACAGCTGATGTTTCCAGAAATTAAGTGAAGCTGCTGCGGCGAAG 3063
Db 3030 TTGAGAGACCGGACCTTACCGACCATTTCCAGAAATTAAGGCGAAGCTGTGTGGCCAG 3089
Qy 3064 GCTTTAGACGAGTCTTGGCCACCGGCAATTTGTCTGACGAGACAGCAGTGGAGAG 3123
Db 3090 GCTTTGATGCGGATCTGAATAATGACGAGCATGACATGACCACTGAACATGAAACT 3149
Qy 3124 TTGACACCATTAAGCATGACAGAGCTTCACTGAAATGCACTGAATCTTTTGC 3183
Db 3150 GTGATTAATTTGAAACCGAACAAGCTCATCAGCAGAGATGATTAACCAACTATGC 3209
Qy 3184 ACCAGGTTCTTGAAGTGAACCTGACAGTGGGTATTTCCGCTCCATACCTGACACT 3243
Db 3210 GTGAGTCTTGTGACCTGACCTGACCTCGGCTCATTTTGTGACCCACCTGTTCCGTTA 3269
Qy 3244 ACTTACAGGATACAGACTGAGGATTAATCGCAGAGGAAAGAAATGATGAGCTTAATGA 3303
Db 3270 TCCATTAAGAAATTAATCACTGAGATTAATCCCTGCTGCTTAACATTAAGGCTGAATAA 3329
Qy 3304 GAGGTAGCAAGAGTGTCAACGCGATATCTGTGATACAAAGCGGTTGACACAGGC 3363
Db 3330 GAAAGTGTCCGCTAGCTCTCCGCAAGGTAACCAAGACTGCTCGAGAGTGTCTACTGA 3389
Qy 3364 AGGAGTGTATTAAGAAATTAATCACTCAAGACTCTCTCAACAAATTAATGTGTT 3423
Db 3390 AGAGTATATGACATGAACACGCGACACTGTGCAATTAATGATCGCGATTAATTAATA 3449
Qy 3424 CCATTAATTCGCGGTTGCCCACTGTTGATCTGTTGACCAAAAGACAGGGTCAACT 3483
Db 3450 CCAAGTAACAAAGACTGCTCATAGCTTAATCTTCCACATTAATGAACACCAAGAGT 3509
Qy 3484 GATCAACGCGATTCATCTTAAGATGAAGGCAATTTGTTGTGTGATGCGGATCTT 3543
Db 3510 GACTTTCTTCAATTCGTGAGCAAACTGAAGGCGAGAACTGTCTGTGTGTGCGGGAGAG 3569

QY 3544 ATCAGCATTCAGGAGGAGAAAGTAGAGTCATGGGTCATTTGCCCATTAATACCATGAGG 3603
 Db 3570 TTGTCCTCCAGGCAAAACGGTTGATCGTTTTCACAGCCGCTGAGGCTTACCTTACGA 3629
 QY 3604 TGTGATCTCGATTGGGAAATACCTAGCCATGTCCGTAATATGACATTTATCTTTGTCAAT 3663
 Db 3630 GCTCGGCTGAGATTAGGATTCCTCCAGGTGCGTCCCAAAATACGACATGTATTATTAAT 3689
 QY 3664 GTTAGAGACCCCGTACAGAAACCATCACTACCAACAGTCCGAGATCAGCTATCCACAC 3723
 Db 3690 GTGAGGACTCCATATTAATACCATCATTTATGAGAGTGAAGACCAAGCCATTAAAGTT 3749
 QY 3724 AGCATGTCAAGCTTAAGGCTGTCCACCACTGAACATGCGGGAACATGTGGCTATA 3783
 Db 3750 AGTATGTTGACCAAGAAAGCTTGTCTGATCTGAATCCGCGGAACCTGCTGACGATA 3809
 QY 3784 GGGTATGGGCTTGTCTGATCCGCAACCGAATATCATCTGCGGTGACGCGTCAATT 3843
 Db 3810 GGTATGTTACGCTGACAGGGCCAGGAAAGCATTTATGTGCTATAGCCGGCAGTTC 3869
 QY 3844 AGGTTTACCCGTGTCTGTACGCTTAAGAAACCTGCCAAAATACTGAGGTTCTTGTG 3903
 Db 3870 AAGTTCCTCCGGGTATGCAACCCAAATCTCACTTGAAGAGACAGAGTTCTGTTGTA 3929
 QY 3904 TTCTTCGCAAGGACCAACGCGCAACACACATGACCGACAGACTCGGTGATAGTCTT 3963
 Db 3930 TTCACTTGGGTACGATCCGCAAGCCCGTACGCAATCTTACAAAGCTTTCACTACCTTG 3989
 QY 3964 GACAACATCTATCAAGGCTCAACAGGTACGAGGCAAGGAGAGCTCCAGCGTACAGATG 4023
 Db 3990 ACCAACTATCTATACAGGTTCCAGACTCCAGAACCGGATGCGACCTCATATCATGTG 4049
 QY 4024 ATCAGAGTGAATTAGCAAGAGGCGTGAACCAACTATCTGTTATGCTGCTATAGCAAA 4083
 Db 4050 GTGCGAGGGGACATTGCGCACCGCCACCGAAGAGTGAATCATTAATGCTGCTAACAGCAA 4109
 QY 4084 GGTCAACCAAGGTTCCGAGGTGCGGTGACCTGTATCCGAAATAGCGCGGCTTTTGAT 4143
 Db 4110 GGACAACTCTGCGGAGGGGTGTGCGAGCGCTGTATGAAGAAATTCGCGAAAGCTTGAT 4169
 QY 4144 AGACAGCCAAATAGTGTGCGGACGCTAGACTTGTGAAGACGAAAC---GCTCATCATTA 4200
 Db 4170 TTACAGCCGATCGAAGTGGGAAAGCGGACTGTGTAAGGTGACGTAACCATATCATTT 4229
 QY 4201 CATCTGTAGAGACCCATTTTCTTAAGATGCCGGAACCGAGGGCGACCTTAAGTCGCA 4260
 Db 4230 CATGCCGTAGGGCCCACTTCACAAAGTTTCGAAAGTTGAAGGTGACAAACAGTTGCA 4289
 QY 4261 GCTGCTACATGAGCATAGCGTCATGCTCAACGCTGACGCGATTAACAAAATATCAAGTA 4320
 Db 4290 GAGGCTTATGATCGTCAATGCTTAAATTTGTCAACGATTAACATTTACAGTCACTAGCCATT 4349
 QY 4321 CCGCTACTGTCAACCGGATCTATTTCTGTGCAAGAGATGAGTGAATGCAATTCATTTGAT 4380
 Db 4350 CCACCTGTTGTCACCGGATCTTTTCGCGAACAAGATGACATTAACCATTCATTTGAC 4409
 QY 4381 CACCTGTTCACTGCTTTTGAACACTACAGATCCGATGTCACCATATATTTGTTGGATTA 4440
 Db 4410 CATTTGTGACAGGTTTGAACACACCATGATGACATGATGACATATCTGACGAGGACAG 4469
 QY 4441 CAATGGGAGACAGATATGAGAGGCAATTCACGCAAAAGAGTTCGAAATTTCTGGAT 4500
 Db 4470 AAATGGAAATGACTCTCAGAGAGCACTGCTGAGAGAGAGCAAGTGAAGAAATATGT 4529
 QY 4501 GATGACAAGCC-----AGTAGACATTTGCTTTGATCAAGGCTCCACCAAC 4545
 Db 4530 ATATCCGACGACTCTTGGGTGACAGAAACCGGATGACAGAGCTGTGAGGCTACATCCGAG 4589
 QY 4546 AGCTCTTTGGAGGACAGACAGTTTACTCCCTCAATGAGGCAAGTTTATTCATACCTG 4605
 Db 4590 AGTTCTTTGGCTGGAAGGAGGCTACAGCAACAGCGATGCAAGACTTTCTCATATTTTG 4649
 QY 4606 GAAGGTACAGATTCAATCAGACGCGCAAGAGCAATTCGCAAAATTCATGCAATGTGGCC 4665

Db 4650 GAAGGAGCAAAATTTGACAGGCGGCCAAGATATGACAGAAATTAATGCAATGTGCCCC 4709
 QY 4666 AACCAATCTAGAGCTATATGACAGATTTGCTTGTACATCTCTGGGGAGATATGTCCAGC 4725
 Db 4710 GTTGCAACGAGGCGCAATGAGCAAGTATGATATATCTTGTGTAACCATATGAGCAGC 4769
 QY 4726 ATCCGCTCCAAATGCGCAGTAGAGAGTCAAGAGCGCTGTGCTCACTCACACTTCCA 4785
 Db 4770 ATTAGTCAAAATGCCCCGTCGAGAGTCCGAAAGCTTCACACCACTTACAGCTTCTT 4829
 QY 4786 TGCCTGTATATTAACGCTATAGCGGCTGACGCGGTATACAGGTTGCGCTTCCGAAAGAA 4845
 Db 4830 TGCTGTGATCATCATGCTATGATCCAGAAAGAGTACAAAGCCTTAAAGCCTCAGTCCA 4889
 QY 4846 GAACAGTCCCGCTATGCTATCATCTTCTGTTGCCAAGTACAGGATCAACAGCGGTGACG 4905
 Db 4890 GAACAAATTAATCTGTGCTCATCTTTCATTTGCCAAGTATGAAATCACTGTGTGTCAG 4949
 QY 4906 AAGCTAAGTACAGCAAAACGATCCTGTTTTCAGGCGCTGATACACCGGCTGTACACCCC 4965
 Db 4950 AAGATCCAGTCTCCACAGCTATATCTGTCTACCTAAGTGTCTGCTATCATTCACA 5009
 QY 4966 AGGAAGTAC 4974
 Db 5010 CGAAAGTAC 5018
 RESULT 8
 AF004458
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AF004458 11420 bp RNA linear VRL 12-JUL-1999
 AF004458
 AF004458.2 GI:5442458
 Venezuelan equine encephalitis virus
 Venezuelan equine encephalitis virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 Alphavirus; VEEV complex.
 1 (bases 8366 to 9182)
 Powers,A.M., Oberste,M.S., Brault,A.C., Rico-Hesse,R.,
 Schmutz,S.M., Smith,J.F., Kang,W., Sweeney,W.P. and Weaver,S.C.
 Repeated emergence of epidemic/epizootic Venezuelan equine
 encephalitis from a single genotype of enzootic subtype ID virus
 J. Virol. 71 (9), 6697-6705 (1997)
 97404681
 9261393
 2 (bases 1 to 11420)
 Wang,E., Barrera,R., Boshell,J., Ferro,C., Freier,J.E.,
 Navarro,J.C., Salas,R., Vasquez,C. and Weaver,S.C.
 Genetic and phenotypic changes accompanying the emergence of
 epizootic subtype IC Venezuelan equine encephalitis viruses from an
 enzootic subtype ID progenitor
 J. Virol. 73 (5), 4266-4271 (1999)
 99214366
 10196323
 3 (bases 8366 to 9182)
 Powers,A.M., Oberste,M.S., Brault,A.C., Rico-Hesse,R.,
 Schmutz,S.M., Smith,J.F., Kang,W., Sweeney,W.P. and Weaver,S.C.
 Direct Submision
 Submitted (16-MAY-1997) Pathology, Univ. Texas Med. Branch,
 Galveston, TX 77555-0609, USA
 4 (bases 1 to 11420)
 Wang,E. and Weaver,S.C.
 Direct Submision
 Submitted (12-JUL-1999) Pathology, Univ. Texas Med. Branch,
 Galveston, TX 77555-0609, USA
 REMARK
 COMMENT
 FEATURES
 source
 1. 11420
 /organism="Venezuelan equine encephalitis virus"
 /mol_type="genomic RNA"

CDS

```
/strain="66637"
/db_xref="taxon:11036"
45..7502
/note="contains nonstructural proteome 1, 2, 3 and 4;
in-frame termination codon is read through at low
frequency"
/codon_start=1
/translation_except="(pos:5658..5660,aa:OTHER)
/product="nonstructural polyprotein"
/protein_id="AAD4335.1"
/db_xref="GI:5442460"
```

CDS

```
/translation="MEKVVHVIDEEDSPRLALQNSPPEPEFAKOVNTDNDHANAFS
HLASKLIETVDPSTLIDIGSAPARARYSKRHACIPMCACEDPRLYKATLTK
NFKETITDELKMKMELAAMSVDPLETEKMLHDESCREGVAVQVADYADGPT
SLYQANGKVAWYIGDTPFPMKLAGAPSYSTWABETVLTANIGCSGDM
ERSRMSILRKLYLPSNNVLSVSGTITHEKRLSLMSLPSFHLRGONTYRC
ETVSCDGVVVKRIASPLQYKPSGVAATMHRBFLCKTDTLNGRSVPCTY
PATLQDNTGLATDVSDADOKLVLGNORIVNGRTORNTKNTLPLVQAFA
RMAKEYKEDDERPLGLRQLVNGCCWAFRRKITSIKRPTQTIKXNSDHSF
VLPRIGSNTLEIGLTRIRKMLEERKESPLIADVEDYACADAEKAEVELRA
ALPLADVEEPTLEADVDMLOEAGASVETPRGLIYTSYAGDKISYALVSPA
VLKSEKSCIHPLAEQVIVTHSGRKGVAEPYHGVVPEGAHAPDODKISESA
TIYNEREPNARYLHIIATHGALNTDEYKIVPSHDEGLYIDIRKCKYKELV
TGLGTGLVDPPEHEFAESLRTAPAPVQPTIGVGVSGSKSGIKSAVTKDL
VSARKENCAEIIROVKMKGLDVNARTVDSVLNGCHPEVETLIDIAFACAGTLR
ALAIIRPKKAVLCGDPKOCGFENMKCKHFNHEICTQVPHKSLDIDYKNSITVA
TLFDKMRNTNPKETKIEIDTSGTKEODDLITCRGWKOLIDYKNSITVA
ASOGLTRGVAVRYKVNENPLIAPTSNVILTRTEDRIYMKTLADGPKITLTK
YQNTETATIEEMQHEAMHILIERPTPTAFOKANVCNAKALVPLTKADIMTT
EOMTVDFETDKAHSAEIVNOLCVRFGLDSDGLSAPLVSILNNWDSPP
NMYGKNSVVRQLSRYPRLPRAVATGVDMNTGTLNCPBRILVNNRLLPVALV
LHNEHPQSDPSFVSKLGRVTLVVGKLSVPGKTVMLSDQEPATRLADLGI
ASPKYDIVEINVRTPYKTHYQOCEDHAIKLSMLTKACLHNPCTCVSIEGYAD
ASESIGAIROPKESRYCKPSLSETEVLFVPIGDRKATNHPYLSSTIYNT
GSRLEHGCASVHYRGDIATTEGVIINANSGOGGCGGCAIKKPEEPDLP
IEVGARLVKGAAKHIIHAGVGENKVSVEGDQKLAAYESIATVNDNNKSAIP
LSTGIFSGNDRLQIOLANHLTALDITDADVAIICRDKKEMBLEKVAERAVEEI
CISDSSVTEPEDAELVHPKSLAGRGYSTDGTSTPELGTFOHAAIDELINA
MMPVATEANEOVCWYILGESMSIRKCPVESASPTPLPCICIAMPERPURL
KASRPEOITVSSPFLPKYRITGVOKIOCSOBIAPSPVPIHRIKLYVTPREX
OSTEGTPEPAPVNDATRTTPEPIITIEEBEDSISLSDGTHOVLVEADHOPPS
VSSSSWSIIPHASDEVDLSILDTLEGASVSEASASATNSYFASMEHVLPA
TFNRPHPAPRTTTPSLAPRASRTLSVSTPEANVITREGEALHLPASRAVS
SRTLSVNPPEGNRVITREEFAFVAOQSGPDAGAYFSSDTEGGLHQOOSQVTL
SEVLERTELISAPRLDOKEBELRKLLNPTPARSRYOSRSMKALITARRI
LOGGHYKAKGVCYCTTHLPVLYSSVNRABSPRYAECANMKENPVTYAST
CTIPEYDALDMDVASCCLDJTASCPKAKLSFPKQHSYLPETTRISAPSAIQNTLON
VLAATKNCNVTQRELPLVDSAAFNECFKRYACNNEWYETKENDIRTEENVN
YITKLGKKAALPAKTHNLMLODIPMDRFVMDLKRDKVTPGTGTHEERPKYQVIO
AADPLATAYLCGIHRELVRNLNAVLPYHILFDMGADPAILIAEHQOPDQVLETD
IASPDKSEDDAMALTMLIEDLGDVADLTLIEAPEGISHLPTTKRFGAMK
SGMELTLEPVNTVINIVIASRVLRELTNSPCAATIGDNIYKGYKSDLMADRCATL
NMEYKIIDAVYGEKAPFCGGFILDCTIVTGAICRVADLKLFLKGLKGLADDEHDD
RRRLHESHTMNRVGLIPELCKAVESRYETVGTISIYMAMTTLASSVKSYSYLRAG
ITLVG"
7538..11305
/note="contains capsid, E3, E2, 6k and E1 proteins"
/codon_start=1
/product="structural polyprotein"
/protein_id="AAC71187.2"
/db_xref="GI:5442459"
/translation="MFPEQPMYPMQPMPEYRNPFAARRRPFRTDPLAMOVETLS
MANLTFKORRDAPEPEPAKKKREAPKQGGGKGGKKNQKKAATKATPKNVOS
GNKKTKKKPKKROBMKLESDFPIMLEKINGYACVYGKLFPMHVEGKIDND
VLAATKTKKASKYDIEADVPONMRAIDFKTHLEPKQYSWHGADYVTEGRFLVK
GVGAKGDSGRPLIDQGRVAVIILGAVNEGRALSVMMNEKGVATYENBOWS
LVTMCLANVTFCAPQPCICYDRKAPATLAMLNVNNDGSGDELELAIVCPARKR
STEELFKEXKLTLPYMAOPICICAVGSCSPALIAEAKVSDGDEGLRLOTSOYDSS
GNLKGRTWRVYMGHTIEIPLHOVSLRSPCHIVDGHGVLTRCAPGSDITMEFK
NLYTHSGSVPEYKENVGRELXTHPRHGEAOAOVYAHQNRGAAYENHILGSEY
DSLSVLSGSSVYTPPAGTSAIVECEGGTKISTINTAKOSCTKECCRYARLO
NDKMYNSDKLPKAGATLKGKHLVPLLAGKCTVPLAPPEMTITFGRSVSLKHPK
NPYLTITRQLADPEHYTHELISEPARNFTYTEKMEFVMGNHBPKEFWOETAAPGN
```

BASE COUNT		3210 a	2817 c	2306 g	2487 t
ORIGIN					
Query Match	42.5%	Score 2123.8;	DB 14;	Length 11420;	
Best Local Similarity	64.6%	Pred. No. 0;			
Matches 3223;	Conservative	0;	Mismatches 1742;	Indels 24;	Gaps 3;
Oy	10	ACTAATGATCCAAATATGAAAGAAATTCACGTTGACTAGTCAAGCCCGTATGTC	69	HGLPHEVITTHYHRYPMNSTIIGLSICAAITVSVAASATWLFCKSRVSLCTREYRLTENA	
Db	30	AATTAACCTAACCAAAATGAGAAAGTTCACTGACATCGAGAAAGACGCCATTTCTC	89	RMPCLAVLCCAGTARAEITWESLDHIMNNQOMFWIOLILPLAALIVTLRLKCVCC	
Oy	70	AAGTCGTTACAGCGGACGTTTCCCAAAATTTGAATCGAAGCAAGCAGCTCACTGACAAAT	129	VPLPLVAVAGAGATVAAETATMPSOAGSVYTIWNRAGYALPLISITPTKIKLPTVN	
Db	90	AGAGCTTGGCAACGAGACTTCCGCACTTTGAGATGAGAACCAAGCAGCTCATGATAT	149	LEVYTHCYKMGDPAIKCCGSOECTPRVPRBOCKVPTGVYPRMGKVAFCPCENT	
Oy	130	GACCATGCGCATGCGCAGAGCGTTTTCGATGCGCAACAAAGCTCATTTAGAGCGAAGTC	189	OVSRAIYWKSDMCDLADHAEYKATKATTSVQATLITVGBHSIVTTVYNGETPVNENG	
Db	150	GACCATGCTAATGCCAGAGCGTTTTCGATCTGGCTTCAAACTGATGAAACGAGAGTG	209	KLTIAGPSTANTPFDRIKIVQAGEIYNDFEYAGQDAGFDIOSRTVSSDIYANT	
Oy	190	GACCGGGAACCAAGTTATCTTGGACATTTGGAAGTGGCGCGCTCAGACATGCAATTC	249	NLVQIRPKAGAIIHVPTQAPSGFQWCKDAPSLKFTAPFCEIYTNPIRANCVGS	
Db	210	GACCATCGCAGACGATCTTGACATTTGGAAGTGGCGCGCGCAAGATGATTTCTAAG	269	ILPLAFIDALFTVSEPTLSAECTNECVSGSDFGIATVYKASGSKCAVHP	
Oy	250	CACCGCTATCTTGTATCTGCCCTATGATAGCGCTGAAGACCCGGAACAGACTCAACAGG	309	SGTATLAEVLETCQGSVTIHFTSTANHPEFRLOICFSYVTCGDCHPDKDHIYTHP	
Db	270	CATAGATATCTTCATCTGCTCGATGAGATGTGCGGAAGATCCGGAACAGATTTGACAG	329	OYHOTEPAANSKRAMTMLSILGGSANVIIIGIVLATIYAMVLTNOKN"	
Oy	310	TATGCAAGAAAGACTTAAGAAAGT-----GACATTACCGACCAAGAACATGAGCCCTTAAG	363		
Db	330	TATGCACTAAGCTGAGAAAGAAATTTCAAGAAATTACTGACAAAGAAATTGACAAAGAA	389		
Oy	364	GCGGAGACCTGCTGGAAGTATGTCACACACAGCAGAGCAGAGACTCCATCTCTGTATG	423		
Db	390	ATGAAGGAGCTCGCGCGCGCTCATGAGGAGCCCTGACCTGGAACCTGAGATATGTCGCTC	449		
Oy	424	CACACAGACCCACGCTGTAGTACTTTTGGAAAGTGAAGTATACCAAGATGTGTAGCCA	483		
Db	450	CACGACGACGAGTGTGTGCTGCTACGAAGGGCAAGTGTGTTTACAGAGATGTATACGCA	509		
Oy	484	GTCATGACGAGCATCAATATCTACACACAGGCGCTTAAGAGAGTTAGGACAAATTACTGG	543		
Db	510	GTTGACGGAACCGCAAGTCTTTATCACCAAGCCAAAGGAGATTAGATCGCCTACTGG	569		
Oy	544	ATAGGCTTTGACACGACCCCTTTTATATGACAAAAACATGAGGCTTCTACCTTACTTAC	603		
Db	570	ATAGGCTTTGACACACCCCTTTTATATGTTAAAGACTTGGCTGAGAGCATATTCATCATAC	629		
Oy	604	AAACAGAACTGGGCTGACGAGAGATTTGGAAGCAGTAACTTTGGCTCGGTAACTCA	663		
Db	630	TCTACCAACTGGGCGGACGAAACCGTGTAAACGCGCTGTAACTAGGCTTATGACAGTTC	689		
Oy	664	GATCTTACGAGACAGAGGCTTGAAGAACTCTCAATCCCTTAGAAGAAAGAGCTCCAACT	723		
Db	690	GACGTCATGAGCGGTACACTGAGGAGTGTCCATTTCTTAGAAGAAAGATTTGAAACCA	749		
Oy	724	ACTAATAAGTCAATATTTCTCGTTGGTGTTCACAATTTACACAGAAATAGATCACTGTTA	783		
Db	750	TCCAAATAATGCTCTATTTCTGTTGGCTGACCATCTTACACGAGAAAGGAGCTTACTG	809		
Oy	784	CGTAGCTGGCATCTTCCAAAGCTGTTCATTGAAAGAAAGTAACTTACAGGTTAGA	843		
Db	810	AGGAGCTGGGACCTGCGCTGTATTTCACTTACGTGGCAAGCAAAATTACACATGTCTGG	869		

844 TGTGGACCATTTGTGACGTGTGAAGGGTACATCAAAAAAGATTAACGATCAGCCAGCA 903
870 TGTGAGACTATAGTATAGTTGCGACGGGTACGTCTTAAAGAAATAGCTATCAGTCCAGGC 929
904 CTATACGGTAAAGTTGAGAACTTGGCGTCCAATATGATCGAGGGTCTTTGAGTTGC 963
930 CTGATAGGGAAGCCATCAGGCTATGTGCGACGATGACCGGAGGATTTCTGTGCTGC 989
964 AAATGTCAGATAGTACGTCGCGCGGAGAGGGTTCTTTGCTGTGTAGTACATGTAACA 1023
990 AAATGTGACAGCATTGAAACGGGAGAGGGTCTCTTTTCCGTGTGACGATATGTCCCA 1049
1024 GCCACATTTGCGATCAGATGACAGGATTTCTGGCAACTGACGTTAGTGTGATGAGCA 1083
1050 GCTACATTTGTGTGACCAATATACCTGGCACTATGGCAACAGATGTCAGTGGCGAGCG 1109
1084 CAAAACTATTTGTTGGGCTCAACCAAGATTTGCTCAATGATGAGACGCAAAAGAAAT 1143
1110 CAAAACTGCTGTTGGGCTCAACCAAGGCAATGCTCAACGGTCCGCAACCAAGAAAC 1169
1144 ACTAACCAATGACGAATCTATTTACAGTGTGCGCCGAGGCTTTTCCAGTGGGCG 1203
1170 ACCAATACCATGAAGAAATTAATCTTTTGCCTGATGTGCGCCAGCAATTTGCTAGGTCGA 1229
1204 CGTGAACATCGTGCCTGCACTTGGACGACGAAAGAACTAGGGGTGCGGAGCGCACTCTT 1263
1230 AAAGAAATTAAGGAAGATCAAGAAAGATGAGAGGCCACTAGGACTACAGATATGACATTA 1289
1264 ACTATGGGCTGCTGCTGGGCTTTCAAGACCCAGAAATACATCCATCTACAGAAACCTT 1323
1290 GTATATGGGGTCTGCTGGGCTTTTAGAAGGACAAAGATTAATCTATTTATAGCGCCA 1349
1324 GGTATCGCAAAATTAAGAAAGTACCTGCCCTCTTTGATCTATTTGATTTCCAGGCTT 1383
1350 GATATCCCAAAACATATCAAAAGTAAACAGCAATTTTCACTATTTGTGCTGCCAGATA 1409
1384 ACCAGCCACGGGCTCGATATGGGCTTCCGCGTGAAGCTCAAGCTGCTGTAACCAACT 1443
1410 GGCAGTAAACATATGGAGATCGGGCTGAGAAACAAATCAGAGAAATGCTGAAAGAGCGC 1469
1444 GTCAAAACCCGACCGGCTATTACAAATGGCCGATGTGAGACATCTGCGTGTTCACAGCA 1503
1470 AAGAGCCGTCACCTCTCATTAATCTGCGAGGACGTACAAAGGCTAAGTGTGCGACCGAT 1529
1504 GAACCTGAAGAAAGTGTGCGACGCGAAAGATCAAGAAAGCCCTGCCACCTTGTCTCCT 1563
1530 GAGGCTAAGAGGTGTGAGAGCCGAGAGGTTGCGCTGCACTACCACTTTTGGCAGCT 1589
1564 GAAATAGAAAAAGAACCGTATAGAGGACAGAAATGACCTCATTTATGCAAGAGCAGAGCA 1623
1590 GATGTTAGAGAGCCCATCTGTGAAAGCCGATGTTGATGCTTACAAAGAGGCTGGGCGC 1649
1624 GTATGCGTGAAGACCAACGAGACATCAAGGGTGAAGATTAACCCAGGCGAAAGAG 1683
1650 GGCTCAGTGAAGACCTCGTGGCTTGAATAAGGTTACCACTATGCGCGGAGAGCAAG 1709
1684 ATTGGGCTTACGCTATATCTTTCAACCCAGCGGATTTGAATATGTAAGAAATCGCGCTGT 1743
1710 ATCGGCTCTTACGAGTGTCTTTCTCAAGGCTGTACTCAAGAGTGAAGAAACTATCTTGC 1769
1744 ATCCACCATTTGGGCGAAACAGATCTGGTATATGACTCAAAAGGTAGGGCAGGGAGATAC 1803
1770 ATTACCTCTCTGCTGAACAAAGTCAATATGATTAACACACTCTGGCCGAAAGAGGCTTAT 1829
1804 AAATGCGACCATACCAAGTAAAGTCAATTTGATCCAGAAAGGAGCGCGGCTCTGTTCAA 1863
1830 GCCGAGAACCTTACCATGAAAGTATGTTGTCAGAGGACATGCAATACCCGTCAG 1889
1864 GACTTCCAGGCAATGATGAGAGCGCTACGATCGTTTTCACAGAGGAGATTTGTAAC 1923
1890 GACTTTCAAGCTCTGAGTGAAGGTGCCACATTTGTTCACAGAACGAGATTTGTAAC 1949
1924 AGATACCTGACACATGCAATCAACGAGGAGCGTAAACATGACGAAGAGTATAT 1983

1950 AGGATCTGACCATATTTGCGACACATGAGAGGCGGTGAACACAGATGAAGATATTAAC 2009
1984 AAGACTGAAGACTCAGACACAGACTCAGAAATAGCTCTTGCATATATGACGACGAAAG 2043
2010 AAATATGTCAAGCCACAGCCGACAGACGGGGAATACCTGTATATACATCGACAGAAACAG 2069
2044 TGTATTAGCGAAGAGACGAGGCTCCCTTGTGCTTAACCGGTATCTGTATATCCACA 2103
2070 TGGCTCAAGAAAGATCTAGTCACTGGGCTTAGGGCTTACAGGCGACGTGTGATCTCTCC 2129
2104 TTTACAGAGTTTGCATACGAGACTCTCAAGACGACACAGACACCTCAAAAGTCCCA 2163
2130 TTCATGAATTTGCGCTACGAGAGTGTAGAACACGACCGGCGCTCTTACCAAGTACCA 2189
2164 ACCATGAGACTATATGAGATGTCAGGTTCCAGTTCAGTAAATCTGGAATCAATCAAAAGCGCTGTG 2223
2190 ACCATAGGGGTATATGCGCGTGGCTTCCAGGAAAGTGTGCATCTTAAGACGAGTC 2249
2224 ACTAAGAAAGATCTGGTGTGATGTCGAGAGAGAAAGAAACTGCGCAGAAATCATACAGGAT 2283
2250 ACCAAAAAGATTTGGTGTGAGCGCCAGAAAGAAAGTCCGACAGATATTAAGGAC 2309
2284 GTAAGAGATGAGACGTATGATGTTGCTGTAGGACTGTGCAATTCAGTCTTCAAT 2343
2310 GTCAAGAAATTAAGAAAGGCTGGAAGTCAATGCGAAGCTGTGACCTAGTCTTGAAT 2369
2344 GGGGTTAAGACCCCGTTAACTCTGTACATTTGATGAGGCAATTTGCTGCTGCTGCAAGG 2403
2370 GGAATGCAAAACCCCGTATGAGACCTGTATATTAAGCAAGCTTTTGTGTATGACGGC 2429
2404 ACGGTGTCGACATGATGTCATGTCATGCAACCTAAGAAAGTGTGATTTGTCGCGGAGCCA 2463
2430 ACTCTACAGCGCTCATAGCATATTAAGCTTAAAGGCAAGTGTCTTGGGGATCCA 2489
2464 AAACATGCGGCTCTTTTAAATGATATGCTGCTGAAAGTACATTTTAACTATGACATATGC 2523
2490 AAACAGTGGGCTTTTCAATATGATGCTGCTGAAAGTGAATTTTAAACAGAGATTTGC 2549
2524 ACTGAATGATCACTAATAAGCATCTCTAGAGGTGCAACAGACTGATACCGCATGTGC 2583
2550 ACGAGGCTCTTCAACAAAGCATCTCTGCTGTGCACTTAATCTGTGACTTTGGTGTGC 2609
2584 TCCAGGCTCTTACAGCAAGCGAATGAAGACGGTTAACCCATGTGCTGTAATAATCATC 2643
2610 TCAACCTGTTTTCAGCAAGAGATGAAGAACAGACACCCGAAAGAGATTAAGATTGAG 2669
2644 ATAGATTCACAGAGGACCAACAAAGCCGACAAAGATGATCTGATTTACCTGTTTCA 2703
2670 ATTGACACTACTGGCAGATTAACCTGACAGAGACGATCTCATTTCTCACTGTTTCAGA 2729
2704 GGAATGGTGAACAGCTACAGATTTGACTAATAAATCAAGAAATCAATGATGCGGCTGCA 2763
2730 GGGTGGTGAAGCAAGTTGAATGATTAACAAAGCAACAAATAAGACGAGCGTCC 2789
2764 TCCGAAGACTTACGCGGAAAGCGTTATGCTGTGCTGATGATCAAGAGTCAAGAGATCCA 2823
2790 TCTCAAGGCTTACCCGTTAAAGGCTGTATGCGGTTGGTATGCAAGGTGAATGAATCCC 2849
2824 CTCTACTCGAGACTTGTGACAGTGAACGTGTTACTTACACGACACAGAAAAACGCAAT 2883
2850 CTGTAGCCACCACTCAGAAACATGTGAACGTCTTACTGACCGGACGAGGACCGATATC 2909
2884 GTCTGAAGACGCTAGCTGTGATCTCTGATTAAGACACTTACAGTAAATATCCCGG 2943
2910 GTGTGAAAAACATAGCGCGGCGATCCATGATTAACACTGACGCGCAAGTATCCGCG 2969
2944 GATTACGGCTTCATTTGAGACAGCTGACGAGGGAACAGACGCAATTAAGGACGCGCTT 3003
2970 AATTTCATCAGCATATGAGATGCAAGAGCATAGAGGCTCATATGAGCACATC 3029
3004 CTGTATAGCCGACAGCTGATGTGTTCCAGATTAAGGTGAACGTCTGTGGCGAG 3063

Dh 3030 TTGAGAGACCGGACCTTACCGACGATTCAGAAATTAAGCGAAGCTGTGTGGCCAAAG 3089
Qy 3064 GCTTTAGAGCCAGTCTTGGCCACGCGCAACATTTGTGCTGACGAGACAGCAGTGGGAGACG 3123
Dh 3090 GCTTTGTGTCCGGTACTGAAATCTGACGGCATATGACATACCACTGAAATGAAACACT 3149
Qy 3124 TTGCAACCAATTCAGACATGACAGAGGATCTACCTGAAATGCACTGAACCTTTTGGC 3183
Dh 3150 GTGAGATTATTTGAAACGGAACAAAGCTCACTCAGCAGAGATAGTATGAAACCAACTATGC 3209
Qy 3184 ACCAGGTTCTTTGGAGTACACTGACAGTGGGTTATTTTCCGCTCTTACCTGCGACTT 3243
Dh 3210 GTGAGGTTCTTGGAGTACACTGACAGTGGGTTATTTTCCGCTCTTACCTGCGACTT 3269
Qy 3244 ACTTACAGGATCAGACCTGGATTAATCTGCGAGGGAAGAACATGATATGGCTTATAGA 3303
Dh 3270 TCCATTAGGAATTAATCACTGGGATTAATTTCCCGTCCCTTACATATGAGGCTGAAATAA 3329
Qy 3304 GAGGTAGCAAGAGATTTGTCAAGGCGATATCCGTGCATCACAAAAACGGTTGACACAGGC 3363
Dh 3330 GAAAGTGTCCGTCAGCTCTCCGACAGGTATCCACGACTGCTCGAGCAGTTGCTACTGGA 3389
Qy 3364 AGGAGTGTGATATTAAGAAATTAATCAATCAAGACTACTCTCAACATTAATGTGTT 3423
Dh 3390 AGAGTCTATGACATGAACACCGGCACTGTGCAATTAATGATCCGCGCATTAATCTAGTA 3449
Qy 3424 CCATTAATATCGCGGTTGCCCCCACTGTTGATCGTTGACCAAAAGACAGGATCAACT 3483
Dh 3450 CCAAGTAAACAGAACTGCTCTATAGTCTTACCTCACCATTAATGAACCCACAGAGT 3509
Qy 3484 GATCAGACGAGATTCCTATTAAGATGAAGGCAAACTGTGTGATTCGCGCATCT 3543
Dh 3510 GACTTTTCTGTGCTGTGAGCAAACTGAAGGCAAGACTGTGTGTGTGGGAGAAAG 3569
Qy 3544 ATCAGCATTTCCAGGAAAGAAATGAATGATCATGGTCCATTGCCCATTAATCAATCAGG 3603
Dh 3570 TTGTCCGTCACGAGAAACGTTGATCTGTGTGACACAGCCTGAGGCTACTTTCAGA 3629
Qy 3604 TGTGATCTGATTTGGGAAATACCTAGCCATGTGCGTAATATGACATTAATCTTGTCAAT 3663
Dh 3630 GCTCGGCTGATTTAGTTAGTATCCAGGTGCGGTCGCAAAATGACATGATTTATTAAT 3689
Qy 3664 GTTAGACCCCTGACAGAAACCATCACTAACAGTGGAGATACAGCTATCCACAC 3723
Dh 3690 GTGAGGATCCATTAATTAATACATTAATGACAGTGTGAAGACACGCGCATTAAGCTT 3749
Qy 3724 AGCATGCTTAAGCTGTAAGGCTGTCCACCACTGAAACATGCGGAAATGTGTGCTATA 3783
Dh 3750 AGTATGTTGACCAAGAAAGCTTGTGCTGAAATCCCGCGGAAACCTGCTCAGCATTA 3809
Qy 3784 GGGTATGGGCTTGTGATCGCGCAACCGGAATATCATCTGCGGTGGCAAGCTCATTT 3843
Dh 3810 GGTATGTTAGCTGACAGGCGCAGCGCAAGCAATTAATGTTGTGTAATGCGGCAAGTTC 3869
Qy 3844 AGGTTTACCCGCTGTCTGACGCTTAAGAAACCTGCGGAAATATGAGGTTCTTCTCGT 3903
Dh 3870 AAGTTCTCCCGGATGCAAAACGAAATCTTCACTTGAAGAGACAGAAATCTGTTTGTGA 3929
Qy 3904 TTCTTGGCAAGACAGCAAGGCAACCACTACATGACAGACAGTCCGTTGTAGTGTCTT 3963
Dh 3930 TTCTATTTGGGTACGATCGCAAGGCGCGTACGCAATCTTACAACTTTCATCTACCTTG 3989
Qy 3964 GACCAATCTTATCAAGGCTCAACCAAGTACAGAGGCAAGGAGGCTCCAGCTTACAGATG 4023
Dh 3990 ACCAAATCTTATACAGGTTCCAGACTCCACGAAGCGGATGTGCAACCTCATATCATGTG 4049
Qy 4024 ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCTGTAATGCTGCTAATAGCAA 4083
Dh 4050 GTGCGAGGGGACATTGCAACGCGCAACGAGAGATGATATAAATGCTGCTTAACAGCAA 4109
Qy 4084 GGTCAACCAAGTTCCGAGTGTGCGGTGACCTGTACCGAAATGCGCGGCTGCTTTGAT 4143
Dh 4110 GGACAACTGTGCGAGGGGTGTGCGAGCGCTGTATAAGAAATTTCCCGAAAGCTTGTGAT 4169

Qy 4144 AGACAGCAATAGTGTGCGGACGCGCTAGACTTGTGAAGACAGAAC---GCTCATCTATA 4200
Dh 4170 TTACAGCCGATCGAAGTGGAAAAAGCGGACTGTGTCMAAAGTGTGAGCTAATATCATTT 4229
Qy 4201 CATGCTGTAGGACCCAAATTTTCTAAGATGCCGGAACCGAGGCGGACCTTAAGCTCGCA 4260
Dh 4230 CATGCCGTAGGGCCCACTTCAACAAAGTTTCGGAAGTTGAAGGTGACAAACAGTTGGCA 4289
Qy 4261 GCTGCTACATAGAGATAGCGTCCATGCTCAACGCTGAGCGGATTAACAAAAATATCAGTA 4320
Dh 4290 GAGGCTTATGAGTCAATGCTGCTAAATTTGTCAACGATTAACATTAAGTCACTAGTACGAT 4349
Qy 4321 CCGTACTGTCAACCGGATCTATCTGTGGCAAAAGTCAGTATGATCAATCATTTGAT 4380
Dh 4350 CCACGTGTGTCAACCGGATCTTTTCCGGAACAAAGATGCACTAACCAATCATTTGATAC 4409
Qy 4381 CACTGTTCACCTTTTGTGACACATACGATGCGGATGTCACCATTAATTTGCTTGATTA 4440
Dh 4410 CATTGTGTGACAGCTTTTGAACCACTGATGAGATGTAAGCATTAATCTGACGAGACAAG 4469
Qy 4441 CAATGGAGACCAAGATTAATGAGGCACTTACCGCAAAAGAGCTGGAATTTCTGAT 4500
Dh 4470 AAATGGAAATGACTCTCAAGGAAGCACTGCTAGAGAAAGCACTGAGAGATATGT 4529
Qy 4501 GATGCAAGGCC-----AGTAGCATTTGACTTGTGAGGATGCCAACCAAC 4545
Dh 4530 ATATCCGACGACTCTTGGGTACAGAACCCGATGAGAGCTGTGTAGGATCATCTGAG 4589
Qy 4546 AGCTTTTGGAGGAGACACAGTTACTCCGTATAGAGGCAAGTTGATTTATCATCTG 4605
Dh 4590 AGTTCTTTGGGTGGAAGAGAGGCTTACAGCAACAGCATGGAAGACTTTCATATTTTG 4649
Qy 4606 GAAGTATACGATTCATCAACACCGCAAGACATTTGCCGAATTCATGCAATGTGGCC 4665
Dh 4650 GAAGGACCAATTTCAACAGGCGGCCAAGATATGACGAATTAATGCAATGTGGCC 4709
Qy 4666 AAACAAATGAGGCTTAATGACAGATTTGCTTGTATACCTCGGGGAGAGATATGCCAGC 4725
Dh 4710 GTTGCAACGAGGCGCAATGACAGATATGATATATCTTGTGAAAGCATGAGCAGC 4769
Qy 4726 ATCCGCTTCAATATGCCAGTATGAGAGATCAGAGGCGTGTGCTCCACTACACATTTCA 4785
Dh 4770 ATTAGTGAATATGCGCGTGTGAGAGATGCGGAGCTTCAACACACTTATGACATGTGCT 4829
Qy 4786 TGCCTGTATTAATGCTATGACGCGTGAAGCGGTATACAGTTGCGCTTGGCAAGAA 4845
Dh 4830 TGTCTGTGATTCATGCTATGACTTCCAGAAAGATACAAAGCTTAAGCTTCAAGTCCA 4889
Qy 4846 GAACAGTGGCGGATATGCTATCATTTCTGTTGCGGAAGTACAGGATCAGAGGCGTGAG 4905
Dh 4890 GAACAAATTAATCTGTGTGCTATCTTTCATTTGCGGAAGTATGAAATACATGCTGTGAG 4949
Qy 4906 AAGTACAGTGCAGCAAAACAGTCTGTTTTCAGGCGTGTATGACACCGCTGTACACCCC 4965
Dh 4950 AAGATCCAGTGTCTCCAGCTTATCTGTCTTCACTTAAGTGTGCTGTATCATATCA 5009
Qy 4966 AGGAAGTAC 4974
Dh 5010 CGAAAGTAC 5018

RESULT 9
BEVCOMGEN
LOCUS 11444 bp ss-RNA linear VRL 01-JUN-2001
DEFINITION Venezuelan equine encephalitis virus, complete genome.
ACCESSION L04653 L00931
VERSION L04653.1 GI:290609
KEYWORDS
SOURCE
ORGANISM
Venezuelan equine encephalitis virus
Venezuelan equine encephalitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; VEEV complex.

REFERENCE 1 (bases 1 to 11444)
AUTHORS Kinney,R.M., Tsuchiya,K.R., Sneider,J.M. and Trent,D.W.
TITLE Genetic evidence that epizootic Venezuelan equine encephalitis
(VEE) viruses may have evolved from enzootic VEE subtype I-D virus
JOURNAL Virology 191 (2), 569-580 (1992)
MEDLINE 93079859
PUBMED 1448915
REFERENCE 2 (bases 1 to 11444)
AUTHORS Kinney,R.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-1993) Molecular Virology, Centers for Disease
Control and Prevention, PO Box 2087, Fort Collins, CO 80522, USA
FEATURES
source
1. 11444
/organism="Venezuelan equine encephalitis virus"
/mol_type="genomic RNA"
/strain="P676"
/isolate="mosquitos/Venezuela/1963"
/db_xref="taxon:11036"
/tissue_1tb="puc18 and M13 of Kinney et al."
1. 11444
/function="genomic mRNA of VEE virus as determined from
strain P676"
1. .7526
/gene="NS"
1. .44
/gene="NS"
45. .7526
/note="possible incorporation of arginine, cysteine or
tryptophan at read through of UGA codon"
/codon_start=1
/trans_except=(pos:5682..5684,aa:Arg)
/product="putative nonstructural polypeptide precursor"
/protein_id="AAC19318.1"
/db_xref="GI:3201593"
/translation="MEKHYVDIEDSPFLRALORSPOFEVEAKQVTDNDHNAARAS
HLASKIETEDPSDTLIDIGAPARMSKTHNCICPMKCAEDPRLKYATKLNK
NCKEITDKEDLKKMELAAVMSDDLEETMCLHDDSCREGEVAYQVYADGFT
SLHQAANKGVAVWIGFTTPEWFKLAGAYPSYSTMADETVLTARNIGLCSDDV
ERSRRGMSILRKXYLKPSNNVLFSVGTIYHEKDLRSMHLPFHLRQKONTVC
ETIYSCGYYVKRIAISPGLYGKPSGAATMREGPLCSKVTDLNGSRVSPCTYV
PATLQDMGTGLATDVSADQKLLVGNRIYVNGRTORNTMKRYLIPVAQAARA
RMAKEYEDODERPLGIRDROLVMGCWAFRRHKITSIYKRPDTOTIIVNSDFHS
VLPRIGSNTLEIGRTIRKLEBHKPSPLITNEDIQEKACDAKEVREBELA
ALPPLADFEPTLEADVLDLQAGASVETPGLIKVTSIAGEDIKTSYQVAVSPOA
VLSEKSLCHPLAEQVIVITHSGRGRVAYEPHGVVPEGHAIPVQDQALSESA
TIVYNERFVNRVYLIATHGALNTBEYKTVKPSBDEYLDIDRQCVKKEV
TGDLTGELVDPPEHFAVESLRTRPAPVQVPTIGYGVSGSKGIIKSAVTKDL
VLSAKKENCALIIDVKMKGLDVNARTVDSVLNGCKHPETLYIDEAFACHAGTIR
ALIAIIRPKAVILCGDPKCGFPMNMLKVHFNHEICTQVPHKSIISRCIKTSYVS
TLFYDKMRTTNPKETKIVITDTSSTKQKODLLITCRGVNQLOIDYKNEIMTAA
ASQGLTRGVAVARYKVENPLVAPTSEHVVALLTRREDRLVMTKLAGDPMIKILTAK
YGFNFTATIEMOAENHDAIMRHILERPDPDVFNQKAVCAKALVPLTKAGIDMT
EOMNTVLFETDKAHSAIYLNQCVRFGLDLSGLFSAPTPLSIRNNHWDNSP
NMQVLANKEVNRQLSRRYPOLPRAVATGSRVDMNTGRLRYNDPRINLVNPRHVALY
LHNHEHPOSSFSVSKLKGRTVAVGKTVSVCKKVDMLSDOEPATPRALDGI
DVPRYDVFVNRTPYKYNHIOOCEDEAIKLSMTKACCHLNGGTCVSIIGYAADR
AVESIIGALIAQKPFKSVCKPSKSHETEVLFVIGYDKARTINPKLSTLTNIT
GSRLEHAGCAPSVHVRGDIATATEGVIINANKGQPGGCGALYKKEPESDLOP
IEGSRKALVGAAGHIIHAGVNPKNYSVEGDKOLAEYSAIKIYNDNNYKVALP
LSTGIFSGNKRDLTOSLNHLTLALDITDVAVALYCRDKKEMTLKEAVARREVEI
CISDSSVTEPDEALVYHVKPSIAGKRGYSTGCKTFSYLEGKFPQAQDAEINA
MPVATEANQVCMYIIIGESWSTIRSKPVESBASITPSTPLCLCHAMTPBEVQLE
KASRPEQITVCSFPPLPKYRITGVOKIOCSQPIIFSPVPAYIHPRKVLVETPVEET
PESPAENOSTEGTEQEPALVNDVATRMPEPIIIEEBEESIIILSKGPHQVLOVE
ADIHGSPVSSSSWSIIPHSDPVDLSIIDLTLQASVTSAGVAENYSYREBELALP
ARVPAPRTVFRNPPHAPRTTRPPLAHASASRTSIYSTPPGNRVITREBELALP
SRASRSASRTSIYVSNPGVNRVITREFEAFVAAQORFPAGAYIIFSSDTGCHLOO
KSVQRTVLSVEVLERTELEISYAPRLDQEKELIRKQLNPTPANSRVSQRVENM
KAITARIILOGLHYLNAEGVECYRTHLPRLISSSNRAFSSPKVAEYCSRMLEK
NFPVASYCIIPEYDAYILDVAGASCCTIDFSPAKLRFPKSKSYLEPTIRAVS
AIQNTLQNVLAATKRNQNTVQMBELPLDGAARVVECFKKYACNNEYWEFTKENPIR

LTENNVNVTIKLGPKAALFAKTHNLNMLQDIPMDRFVMDLKRDKVTPGTGTEE
RPKVOVQAADPLATADLGIHRELVARNALVALLNHTLEISAEFDIILIEHFP
GDCVLETDIASPKSDEDDMALATMALITLAEALDITLIEAFSGISLPTPK
PKGAMKSGMFLITFNTYINIVIASRVIRELTGSPCAAFIGDDNIVGVYSDKLM
ADRCATWLNMEKVIIDAVVEKAPYFCGGTIGLVGICRAVDPLKRIFKLGKPLA
VDEHDDRRRALHEESTRNVRGIIPELCKAVESRYETVGTISIIVAMTTLASVKS
FSYLRGAPITLYG"
45. .1649
/gene="NS"
/product="putative nonstructural protein nsP1"
/function="minus strand RNA synthesis; methyltransferase;
guanylyltransferase"
1650. .4031
/gene="NS"
/product="putative nonstructural protein nsP2"
/function="RNA helicase; nonstructural proteinase;
necessary for subgenomic 26S mRNA synthesis"
/function="replication"
4032. .5702
/gene="NS"
/product="putative nonstructural protein nsP3"
/function="replication"
5703. .7520
/gene="NS"
/product="putative nonstructural protein nsP4"
/function="RNA polymerase"
/function="replication"
45. .5684
/gene="NS"
/codon_start=1
/product="nonstructural polypeptide precursor"
/protein_id="AAC19320.1"
/db_xref="GI:3201594"
/translation="MEKHYVDIEDSPFLRALORSPOFEVEAKQVTDNDHNAARAS
HLASKIETEDPSDTLIDIGAPARMSKTHNCICPMKCAEDPRLKYATKLNK
NCKEITDKEDLKKMELAAVMSDDLEETMCLHDDSCREGEVAYQVYADGFT
SLHQAANKGVAVWIGFTTPEWFKLAGAYPSYSTMADETVLTARNIGLCSDDV
ERSRRGMSILRKXYLKPSNNVLFSVGTIYHEKDLRSMHLPFHLRQKONTVC
ETIYSCGYYVKRIAISPGLYGKPSGAATMREGPLCSKVTDLNGSRVSPCTYV
PATLQDMGTGLATDVSADQKLLVGNRIYVNGRTORNTMKRYLIPVAQAARA
RMAKEYEDODERPLGIRDROLVMGCWAFRRHKITSIYKRPDTOTIIVNSDFHS
VLPRIGSNTLEIGRTIRKLEBHKPSPLITNEDIQEKACDAKEVREBELA
ALPPLADFEPTLEADVLDLQAGASVETPGLIKVTSIAGEDIKTSYQVAVSPOA
VLSEKSLCHPLAEQVIVITHSGRGRVAYEPHGVVPEGHAIPVQDQALSESA
TIVYNERFVNRVYLIATHGALNTBEYKTVKPSBDEYLDIDRQCVKKEV
TGDLTGELVDPPEHFAVESLRTRPAPVQVPTIGYGVSGSKGIIKSAVTKDL
VLSAKKENCALIIDVKMKGLDVNARTVDSVLNGCKHPETLYIDEAFACHAGTIR
ALIAIIRPKAVILCGDPKCGFPMNMLKVHFNHEICTQVPHKSIISRCIKTSYVS
TLFYDKMRTTNPKETKIVITDTSSTKQKODLLITCRGVNQLOIDYKNEIMTAA
ASQGLTRGVAVARYKVENPLVAPTSEHVVALLTRREDRLVMTKLAGDPMIKILTAK
YGFNFTATIEMOAENHDAIMRHILERPDPDVFNQKAVCAKALVPLTKAGIDMT
EOMNTVLFETDKAHSAIYLNQCVRFGLDLSGLFSAPTPLSIRNNHWDNSP
NMQVLANKEVNRQLSRRYPOLPRAVATGSRVDMNTGRLRYNDPRINLVNPRHVALY
LHNHEHPOSSFSVSKLKGRTVAVGKTVSVCKKVDMLSDOEPATPRALDGI
DVPRYDVFVNRTPYKYNHIOOCEDEAIKLSMTKACCHLNGGTCVSIIGYAADR
AVESIIGALIAQKPFKSVCKPSKSHETEVLFVIGYDKARTINPKLSTLTNIT
GSRLEHAGCAPSVHVRGDIATATEGVIINANKGQPGGCGALYKKEPESDLOP
IEGSRKALVGAAGHIIHAGVNPKNYSVEGDKOLAEYSAIKIYNDNNYKVALP
LSTGIFSGNKRDLTOSLNHLTLALDITDVAVALYCRDKKEMTLKEAVARREVEI
CISDSSVTEPDEALVYHVKPSIAGKRGYSTGCKTFSYLEGKFPQAQDAEINA
MPVATEANQVCMYIIIGESWSTIRSKPVESBASITPSTPLCLCHAMTPBEVQLE
KASRPEQITVCSFPPLPKYRITGVOKIOCSQPIIFSPVPAYIHPRKVLVETPVEET
PESPAENOSTEGTEQEPALVNDVATRMPEPIIIEEBEESIIILSKGPHQVLOVE
ADIHGSPVSSSSWSIIPHSDPVDLSIIDLTLQASVTSAGVAENYSYREBELALP
ARVPAPRTVFRNPPHAPRTTRPPLAHASASRTSIYSTPPGNRVITREBELALP
SRASRSASRTSIYVSNPGVNRVITREFEAFVAAQORFPAGAYIIFSSDTGCHLOO
7524. .7561
/note="putative"
/function="noncoding segment between nonstructural and
structural genes"
7532. .11444
/gene="26S mRNA"
7532. .11444
mRNA


```
/gene="26S mRNA"
/product="26S mRNA region"
/notes="The structural proteins of the virus are translated as apolipoprotein precursor molecule from an intracellular, 26S mRNA species that is identical to the 3'-one-third portion of the viral genomic mRNA; putative"
/function="intracellular, subgenomic viral mRNA species"
7562..11329
/gene="26S mRNA"
/codon_start=1
/product="structural polyprotein precursor"
/protein_id="AAC19319.1"
/db_xref="GI:290611"
/translation="MFPQPMYPMQPMYRNPFAARRRPMFPRTPDPLAMOVLETRS
MANLFEKORADAPMEPRPAKPKREAPKOKGGGCKKKKKKKKAKKATGPNPAOS
GNKKPKPKPKRQRMWMLKSDTFPIMLBKINGYACVVGKLPFRMHYEGKIDND
VLAALKTKASKTDLLEADVPQNMKADTFKTHEKPKQYISMHGAUVYENGKFTVPK
GVGAKGDSGRPLDNOGRVVAIVLGVNVEGRTALSVMMNEKGVTVRTPBENCOWS
LVTMCLLANTYTPCAEPPICDYRPAETLAML SVNVNPGYDELLEAVCPGRKR
STELFEKYLTRPYMARCICAVGSCSPLEIAVKDGDGVRLDTSOYGLDS
GNLKRTRVMHGTEIEIPLHOVSLHSRPHIVDGHGFLAACPASDITMERFK
GSYHSCSVPEYKPNVGRBLTPHPHGAQACQVYAHNORNGAYENHLPSEV
DSLISLSSSVTVTPVGTSLVLCCKGKTSITINKAKQFSCITKEQCRARLQ
NDKVVNSDKLPKAGATLKGHLVPLADGKCTVPLAPPMITFFGRYSKLHPK
NPYLTTRQLADEPHYHELISEPAVRNFTVEKMEFVWGHPKRFMAOETAPGN
HGLPHEVTHYHYRPMSTILGLSICAIIVTVSAVASTLWLFCKSVSCLTPRYLPNA
RMPLCLAVLCCARTARATWESLDHLMNNNOQMPVLIPLAALIYVTLKVC
VVPFLVAGAGAGAYEHATTMPGAGISYNTIVNRAIYALPISITPTKILPTVN
LEVTCHYKTDSPALIKCGSQECTPFRNDEQKVTGFPFMWGAJCFCEJENT
QVSKAYVMSDDCLADHAAYKANTASVOAFLNITVGHSLVTVYVNGEFPVNGV
KLTPGLSTAMTPEDRKIVQYAGEIYNDDPEYGAQGGADIOSRFSISDLVANT
NLVLRPKAGAIHVEYTOAPSGFEOMKDKPSLKTAPFGCEIYNIRBENCASVS
IPLAFIDPALFTYRSETPTLSAECTINECVSDRGIATVYVSAKSKCAVHP
SGTATKAAAEITPQGSATIHFSANIHPEPRLOICTSYTCKDCKPPDHVTHP
QYHAKTFYAVSKTAMTWLISLGGSAVITIIIGLVATVANYVLTNOKN"
7562..8386
/mat_peptide
/gene="26S mRNA"
/product="core protein"
/function="encloses the genomic mRNA molecule"
/standard_name="capsid protein"
/notes="putative"
8387..8563
/mat_peptide
/gene="26S mRNA"
/product="viral cleavage product"
/standard_name="B3 protein"
Query Match 42.4%; Score 2122.2; DB 14; Length 11444;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3225; Conservative 0; Mismatches 1748; Indels 24; Gaps 3;
```

```
QY 356 CCTTAAGCGCGAGACCTGCTGGAAGTCATGTCAACACAGACGACAGATCCATCTC 415
DB 382 ACAAGAAAATTAAGAGAGTCCGCCCGGTCATGAGACCTGACTGAAAATGAGACTA 441
QY 416 TGTGTAATGACACAGACGCCAGTGTAGTACTTTGGAAGTGTAGCACTATACCAAGATG 475
DB 442 TGTGCTCCAGACGATGATGATGTGCTGCGTACGAGGGGCAAGTGGCTTTTACAGGANTG 501
QY 476 TGTACGAGTCATGACACCGACATCATCTACACACAGGGCCTTAAGAGATTAGACAA 535
DB 502 TATACGAGTGTAGACGACCGCAAGTCTCTATCAACCAAGCAAGGAGTTAGATCG 561
QY 536 TTTATGGATAGGCTTTGACACGACCCCTTTATGTACAATAAATGCGAGTTCTTACC 595
DB 562 CTTACTGATAGGCTTTGACACACCCCTTTATGTAAAGAACTTGCTGAGACATATC 621
QY 596 CTACTTACACACGAACTGGGCTGACGAGAGATATTGGAAGCACTTAACATTGGCCTCG 655
DB 622 CATCATACTTACCAACTGGGCGGACGAAACCGTGTAAACGGCTGTACATAGGCTAT 681
QY 656 GTAACTCAGATCTTCAGAGACGAGGCTTGGAAAACTCTCAATCCTTAGAAGAGAGC 715
DB 682 GCAGCTCCGAGCTATGAGCGGTCACGTAGAGGATGTCCATCTTAGAAGAAATTT 741
QY 716 TCCAACTCACTAATAGATCAATATTTCTCGTTGTTCAACAATCTACACAGAAATAGAT 775
DB 742 TGAACCATCAATAATAGTCTATTTCTGTGTGCTCAGCATCTACACAGAAAGAGG 801
QY 776 CACTGTTCAGTACGCTGCATCTTCCAAACGTTTCCACTTGAAGAAAGTCTAATCTCA 835
DB 802 ACTTACTAGAGAGCTGGACCTGCGCTGTATTTCACTTACGTGGCAAGCAAAATTAACA 861
QY 836 CAGGTAGTGTGGACCACTTGTACGTGTGAAGGATCATCAAAAATTAACATCA 895
DB 862 CATGCTCGTGTGACATATAGTTAGTTGTCGACGGGTACGTGTAAAGAAATAGCTATCA 921
QY 896 GCCCAGACCTAATAGGTAAGTGTAGAACTTGGGCTCCAAATGATTCGAGGTTTCT 955
DB 922 GTCCAGGCTGTATAGGAGAGCTTCAGAGCTATGCTGTACATGACCGCAAGGATTTCT 981
QY 956 TGAATTGCAAGTACAGATACGCTGCGCGGCGAGAGGTTCTTTTGTCTGTGTACGT 1015
DB 982 TGTGCTGAAAGTACAGACATTTGAACGGGAGAGGATCTTTTCCCTGTGACGT 1041
QY 1016 ATGTACCAAGCCACTTTCGATCAATGACAGGATTTTGGCAACTGACGTTAGTGTG 1075
DB 1042 ATGTGCCAGTACATTTGTGTGACCAATGACTGACATCTGGCAACAGATGTCAAGTGGG 1101
QY 1076 ATGACGCAAAAATCTATTTGTTGGGCTCAACCAAGATTTGTCTCAATGTAGACGC 1135
DB 1102 ACGACGGCAAAAATCTGTTGGGCTCAACAGCGATAGTGTCAACGTTGCGCACCC 1161
QY 1136 AAAAATACTAATACATATGATGAGAACTATCTATTACAGTGTGCGCCAGCGTTTCCA 1195
DB 1162 AAAGAAACCAATACATGAAATTAATCTTTTGGCCGTAGTGGCCAGGACATTTGCTA 1221
QY 1196 GGTGGGCGCTGAACATGTCGCGCACTTGAAGCAAGAAAGAACTAGGGGTGCGGAGC 1255
DB 1222 GGTGGGCAAAAGATATTAAGAAATCAAGAAAGATGAGAGGCCACTAGACATTAAGAGATA 1281
QY 1256 GCATCTTCAATATGCGCTGCTGCTGGCTTTCAAGACCCCAAAAATCATCTCATCA 1315
DB 1282 GACAGTTAGTATGGGTGCTGCTGGCTTTTAAAGGCAACAAATTAATCTTATTAATA 1341
QY 1316 AGAAGCTGTGTAACCAAACTATTAAGAAAGTACCTGCGTCTTGAATCTATTTGATTC 1375
DB 1342 AGCGCCAGATACCAACCAATCATCAAGTGAACAGAGATTTTCACTCATTTGCTGCTGC 1401
QY 1376 CAGGCTTACAGGCTCAGGCTGATATGGGCTTCCGCGTAGGCTCAAGCTGCTGCTG 1435
DB 1402 CCAGGATAGGACATTAACACTGAGATCGGGCTGAGAAAGAGATCAGGAAATGCTAG 1461
```

Qy 1436 AACCACTGTCAAAACCGGCAACGGCTATTACATAGCCGATGTAGAGCATCTGGCTGCT 1495
Db 1462 AAGAGCAACAAGGACCCCTCACTCTCATTTACTGGCGAGGACATACAAAGAGCTAAAGTGGC 1521
Qy 1496 TACAGCAAGAACTGAAAGAGTGGCTGACGGGAAAGAGATCAGAAAGCCCTGCCACCTT 1555
Db 1522 CAGCGGATGAGCTTAAGGAAATGGCTGAAAGCCGAGGAGCTGGCGCTGCTTACCACTT 1581
Qy 1556 TGCTCCCGGAAATGAAAGAAAGACCCGTAGAGGAGAGTAGAGCTCACTTATGCAAGAG 1615
Db 1582 TGGAGCTGATTTTGGAGAGCCCACTCTGGAAGCCGATGTGACTTGTATTAACAAGAG 1641
Qy 1616 CAGAGCAGGTAGCGGTGGAGACACCAGAGGACATCAGGGTGAACAATTACCAGCGC 1675
Db 1642 CTGGGGCCGGCTCACTGGAGACCTGTGGCTTGATTAAGATTACCAAGCTATGCCGCG 1701
Qy 1676 AAGAAAGATTGGCTCTTACGCTATCTTCAACCCAGCGCGTATGTAATGTGAAAAAC 1735
Db 1702 AGGACAAAGATGGCTCTTACGCGAGTGTCTTCCACAGGCTGTACTCAAGAGTGAAGAAC 1761
Qy 1736 TGGGCTGTATCCACCCANTGGCGGAAACAAGTACTGTAACTGACTCAACAAGTGGGAG 1795
Db 1762 TATTTGATTCACCTCTCTGCTAACAAGTCATGTATTAACAACACTTGCCGAAAAAG 1821
Qy 1796 GAGAGTACAAAGTGCAGCCATACCAACGTTAAGTCAITGTACAGAAAGGACGGCGTCC 1855
Db 1822 GGGCTTATGCCGTGGAAACCTTACCATGGAAAAAGTATGTGTCCAGAGGAGCATGCATAC 1881
Qy 1856 CTGTTCAAGACTTCCAGGCAATTGAGTGAAGCGCTACGATCGTTTCAAGAGAGGAGT 1915
Db 1882 CCGTCCAGGACTTTCAGAGCTCTGAGTGAAGTGCACATCGTGTACAAACAAGAGAGT 1941
Qy 1916 TCGTAAACAGATACCTGCAGACACATCGCAATCAAACGAGCGGCTAAACACTGACGAAG 1975
Db 1942 TCGTAAACAGTACTGCAACATATTGCCACATGAGAGACCGCTGAACAAGATGAAG 2001
Qy 1976 AGTACTATAGACTGTAAAGACTCAGGACACAGACTCAGAACTAGCTCTTCAATATGACG 2035
Db 2002 AATATTCAAAAAGTCAAGGCCAGCGAGCAAGAGCGGAATACCTGTACACATTCGCA 2061
Qy 2036 CACGAAAGTGTATTAAGCGAAGACGCAAGTCCCTTGTGCTTAAACCGTATCTGGTAG 2095
Db 2062 GGAACCATGTGCTCAGAAAGAAATTAATGCTAGTGGCTTAAGGCTTACAGGCGACGTGGTG 2121
Qy 2096 ATCCACCATTTACAGAGTTTGCATACGAGAGTCTCAAGACAGACACACACTGACA 2155
Db 2122 ATCTCTCTTCCATGAATTTTCTTACGAGAGTCTGAGAACAGTCCGCGCTCTTAAC 2181
Qy 2156 AAGTCCCAACCATGCGAGTCTATGAGAGTCCAGGTTCAAGTAACTTGAATCAATCAAAA 2215
Db 2182 AAGTACCAACCATAGGGGTGTATGGCGTCCGGGGTCAAGCAAGTCTGGCATTTAAAA 2241
Qy 2216 GCGCTGTGACTTAAGAAAGATCTGTTGTAGTGGAGAGAGAAAGCTGGCGAAGATCA 2275
Db 2242 GCGGAGTCAACCAAAAAGATCTGGTGTGAGCGCCAGAAAGAAAGAAAGAACTGCCAGAAATA 2301
Qy 2276 TCAGAGAGTGAAGAGAGTGAAGAGTATGATGTTGTGCTAGGACTGTGATTCAGTGC 2335
Db 2302 TAAAGGAGTCAAGAAATGAAGAGGCTGAGCGTACGATGCAAGAACTGTGAGCTAGTGC 2361
Qy 2336 TTCTAAATGGGGTTAAGCAACCCCGTTAAACACTGTGTACATTGATGAGCAATTTGGCTGCC 2395
Db 2362 TCTTGAATGATGCAAAACCCCGTAAAGACCTGTATATTGACGAACCTTTTGGTTGTC 2421
Qy 2396 ATGAGGAGCGCTGTGCACTGATTTGCTCATGTCAAACTTAAGAAAGTGTATTTGTCG 2455
Db 2422 ATGAGAGGCACTCTCAGAGCGCTCATAGCCATCATTAAGACTTAAAGAGCAAGTGTCTGCG 2481
Qy 2456 GGGAGCCCAAAACANTGGCGCTTTCTTAAACATGATGTGCTGAAAGTAAATTTTAAACATG 2515
Db 2482 GGGAGTCCAAAACAGTGTGGCTTTTCAATATGATGTGCTGAAAGTGAATTTTAAACATG 2541
Qy 2516 ACATATGCACTGAAGTGTACCAATAAAGCATCTCTAGAGAGGTGCACACAGACTGTAAACG 2575

Db 2542 AGATTTSCAGCAGAGTCTTCCACAAAAGCATCTCTGCGCGTGCACATAATCCGTGACTT 2601
Qy 2576 CCATGCTTCCAGCTCTTTTACGACAGGAAATGAAGAGGTTTAAACCCATGTGCTGATA 2635
Db 2602 CCGTGTCTCAACCTGTGTTTACGACAAAAGAGTGAAGAGAGCAAGAACCCAAAAGAGCTA 2661
Qy 2636 AAATCATCATAGATACACAGAGGACCAAGACCGGCAAAAGATGATCTGATTTAACT 2695
Db 2662 AGATTTGATTGACATCACTAGGCAAGTACCAACACGAGAGAGAGAGATCTCATTTCTCACTT 2721
Qy 2696 GTTTACAGAGATGGGTGAAAACGCTACAGATTGACTAACAAAATCAGAAATCATGACTG 2755
Db 2722 GTTTCAAGAGGTGGGTGAAACAGTTGCAAAATAGATTCAAAAGCCAAAGAAATATATGACG 2781
Qy 2756 CCGCTGATCGCAAGACTTAACCGGAAAGCGTTTATGTGTACAGGTAAAGTCAAG 2815
Db 2782 CAGTGTCTCTCAAGGGCTACCCGTAAAGCGGTATGTGCGTCCGTACAAAGGTGAATG 2841
Qy 2816 AGAATCCACTCTACTCGCAGACTTCTGAGACGTAAGCGTGTAACTTACACGACAGAAA 2875
Db 2842 AAAATCCCTGTACGACACCACTCAGAACATGTGAACGTCTACTGACCCGACGGAAG 2901
Qy 2876 AACGATTTCTGGAAGACGCTAGCTGTATCTCTGATTAAGACACTTACAGCTTAAT 2935
Db 2902 ACCGTATCTGTGGAAGAAACACTAGCCGCTGATCCATGATTAATAATCTGACGCGCAAGT 2961
Qy 2936 ATCCCGGGATTTTACCGCTTCAATGAGACAGTGGGACCGCGAAACAGAGCCATTTATG 2995
Db 2962 ATCTGGAACTTCACTGCAACAGTAAAGAAATGGCAAGACATGAAATCAATATGA 3021
Qy 2996 CACGCTTCTGATTAAGCCGACAGACGCTGATGTTCAGAAATAGGTGAACGTCTGCT 3055
Db 3022 GGCACATCTTGAAGAGACCGGACCTTACCGAGCTTCCAAAATTAAGCGGAAGTGTGTT 3081
Qy 3056 GGGCGAAGGCTTTAGAGCCAGTCTTGGCCACGGCCAACTTTGTCTACGAGACAGACT 3115
Db 3082 GGGCCAAAGGCTTTGTGTCCCGTACTGAAGACTGAGGCAATGACATGACCACTGAACAT 3141
Qy 3116 GGGAGAGTTGACCCCATTAAGCATGACAGAGCTACTGACCTGAAGTGGCACTGAAC 3175
Db 3142 GGAACACTGTGATTACTTGAACGGAACAGTCACTGACGAGATGATTTGAAC 3201
Qy 3176 TCTTTGCACAGGTTCTTTTGAAGTACCTGACAGTGGGTTATTTCCGCTCTACCG 3235
Db 3202 AACTATGCGTGAAGTTCTTTTGAAGTGAACCTGAGCTCGGTTCTTATTTTTCGACCACTG 3261
Qy 3236 TCGCACTTACTTACAGGAGTACAGCTGGGATTAATCTGCCAGGAGAAACATGTATGGC 3295
Db 3262 TTCCGTTATCCATTAGGAATATCACTGGGATTAATCCCGTGCCTTAACATGTATCGGCT 3321
Qy 3296 TTATTAAGAGTGAAGAAAGAGTGTCAAGCGCATTCCTGTGATCACAAAAGCGGTTG 3355
Db 3322 TGAATTAAGAGAGTGTCCGACGCTCTCCGACAGTAAACCAACATCTCTCGACCAATTG 3381
Qy 3356 ACAAGCAGAGGAGTGTATTAAGAAATATACATCAAGAGACTACTCTCAACAAATTA 3415
Db 3382 CCACCGAAGAGTGTATGAATGAACACTGGACGCTGCCCAATTAATGATCCGGCATA 3441
Qy 3416 ATGTGTTCCATTAATATGCGCGGTGCCCACTGCTTGTATGTGACCAAAAGGACAGG 3475
Db 3442 ATCTAGTACTGTGAACAGAAAGACTGCTCATGTGTTAGTCTCCACCAATATGAACACC 3501
Qy 3476 GTACAACTGATCAAGGAGTCTCTATCTTAAGAGTGAAGGCGAAATGTGTGTGTATG 3535
Db 3502 CACAGAGTGAATTTTCTTCAATTCGTACGAAATGAAAGGCGAAGACTGTCTTGTGTGTCG 3561
Qy 3536 GCGATCTTATCAGCAATTCAGAGGAAAGAAAGTGAAGTCAATGGGCTCATGGCCCATATA 3595
Db 3562 GGGAGAGTGTCTCGTCCGAGGAAAGAGTCACTGTGTGTGACAGACAGCTGAGGCTA 3621
Qy 3596 CCATCAGGTGTATCTGATTTGGGAATACCTAGCCATGTGCGTAAATATGACATTATCT 3655

Db 3622 CTTTGAAGCTCGGTGATTTAGGTATCCAGGTGACGTCGCCAAATACGACATTGTAT 3681
 Qy 3656 TTGTCAATGTTAGACCCCGTACAGAAACCATCACTACCAAGTGGAGATGACGGTA 3715
 Db 3682 TTATTAAGGTAGAGACTCATATTAATACATCATTTATCAGCAGTGTAGACCGACCA 3741
 Qy 3716 TCACACACAGCATGCTAAAGTGTAAAGCTGTCCACACCTGTAACA CTGGCGAATGTG 3775
 Db 3742 TTAACCTAGTATGTGACCAAGAAAGCTGTCTGCATTGTAATCCCGCGAATCTGG 3801
 Qy 3776 TGGCTATAGGGTATGGGCTGTCTGATCCGCAACCGAATATCTCACTGGCGGTGAC 3835
 Db 3802 TCAGCATAGTATGTATGCTGACAGGGCCAGGAGAGCATCTGTTGTCTAATAGCGC 3861
 Qy 3836 GCTCATTTAGGTTTACCGGTGTCTGTACGCTTAAAGACCTGCGAAATATCTAGGTTT 3895
 Db 3862 GGCAGTTCAAGTTCTCCGGGTATGCAACCGAAATCTCACAATAGAGAGAGATGAC 3921
 Qy 3896 TCTTGTGTCTTTCGGCAAGGACCAAGCGACACACATGACAGGACAGACTCGGTG 3955
 Db 3922 TGTGTATTTATGTTGGGTACGATCCGACGGCCGTACGACATCTTACAGCTTTGAT 3981
 Qy 3956 TAGTCTTGACAACTTATCAAGGTCAACCAAGTACGAGCGAGAGAGCTCCAGCGT 4015
 Db 3982 CTACCTTGACCAACATCTATACAGGTTCAGACTCCAGCAAGCCGATGCGACCTCAT 4041
 Qy 4016 ACAGATGATCAGAGGTGACATTAGCAGAGCGGTGACCAAGCTATCGTTATGCTGCTA 4075
 Db 4042 ATCATGTGTGCGAGGGATATGCGACGCGACCGAAGAGATGATCAATTAATGCTGCTA 4101
 Qy 4076 ATAGCAAGGTCAACCAAGTTCGGAGTGTGCGGTGACCTGATACGAAATGGCGGCTG 4135
 Db 4102 ACAGCAAGGACCACTGCGGAGGGGTGTGCGAGCGCTGTATAGAAATTTCCCGGAAA 4161
 Qy 4136 CTTTGTATGACAGCCATAGCTGTGCGGACGGTGAAGCTGTGAGACGAGAC--GC 4192
 Db 4162 GCTTCGATTTACAGCGCATCGAAGTGAAGAAAGCCGACTGTCAAGGTGACGTTAAAC 4221
 Qy 4193 TCATCATCATGCTGTAGAACCACTTTTCTAAGATGCGGGAACCGAGGGCGGCTTA 4252
 Db 4222 ATATCATTCATGCTGTAGAACCACTTTCAACAAAGTTTGGAAAGTGAAGGGGCAAC 4281
 Qy 4253 AGCTGCAAGTGTCTACATGAGCATAGGCTCCATGCTCAACGCTGAGCGATTAACAAA 4312
 Db 4282 AGTTGGCAGAGCTTATGAGTCCATCGCTAAATTTGTCAACGATTAACAATTAAGTCA 4341
 Qy 4313 TATCAGTACCGTACTGTCAACCGCATCTATTTCTGTGGCAAGATGAGTATGATCAAT 4372
 Db 4342 TAGCATTCACACTGTGTTCACCGGCATCTTTCCGGGAACAAAGATGACATCAACCAAT 4401
 Qy 4373 CATTCATCACCTGTCTCTGCTTCGACACTACGGAATGCCATGCTCAACCATATTTGCT 4432
 Db 4402 CATTTAACCATTTGCTGACAGTTTAAACACACATGATGCAATGTAGCCATATCTCA 4461
 Qy 4433 TGGATTAACAATGAGGACCAAGATATGAGGACCATTCACCGCAAGAAAGAGCGTCAAA 4492
 Db 4462 GGGAAAGAAATGGAAATGACTCTCAAGAGACAGTGGCTGAGAGAGAGAGTGGAGG 4521
 Qy 4493 TTTCGGA-----TGATGACAAAGCAGTACGATGACCTTGTGAGGGTTC 4537
 Db 4522 AGATATGATATCAGACGACTTTCGTGACAGAACCGGATGACAGCTGATGAGGTAC 4581
 Qy 4558 ACCCAAGAGCTTTTGGCAGGCAAGAGTTACTCCGTCAATGAGGCAAGTTGTAT 4597
 Db 4582 ATCCGAAGAGTTTGTGCTGAGAAAGGCTACAGCAAGATGAGCAAGCTTTCT 4641
 Qy 4598 CATACCTGAAGGTACAGATTCATCAGACCGCAAGGACCTTCCGAAATCATGCA 4657
 Db 4642 CATATTTGGAAGGACCAATTTCAACGAGGGCCAGAGATATAGCAAAATTAATGCA 4701
 Qy 4658 TGTGGCCCAAAATCTGAGGCTAATGACAGATTTGCTGTATCTCTGGGGAGAGTA 4717
 Db 4702 TGTGGCCAGTTGCAACGAGGCGCATGAGCAAGTATGATATCTCTCGGTGAAGCA 4761

Qy 4718 TGTCCAGATCGGCTCCAAATGCCAGTACAGAGGTGAGAGCGGTGCTGCTCACTCA 4777
 Db 4762 TGAGCAGATTTAGTTCGAAATGCCCCGTGAGAGTCCGAAGCTTCACACCACTAGCA 4821
 Qy 4778 CACTTCATGCTGTGTATTAACGTATACAGCGGTGAGCGGCTATACAGTTCGCTGTG 4837
 Db 4822 CGTTCCTGCTTGTGATCATGCTATGATCTCCGAAAGAGTACACAGCCCTAAAGCT 4881
 Qy 4838 CGAAGAAAGACAGTTCGCCGTATGCTATCATTTCTTTGCGGAAGTACAGATACAG 4897
 Db 4882 CACGTCCGAACCAATATCTGTGCTCATCTTTTCCATTCGGAAGTATGAAATCACTG 4941
 Qy 4898 GCGTCAGAGCTACAGTACGACCAACAGCTCTGTTTACGAGCGCTGTACACCGGCTG 4957
 Db 4942 GTGTGCAAGATTCAGTGTCTCCAGCTTATACCTTTCTACCGAAGGTGCTGCTACA 5001
 Qy 4958 TACACCCGAGAGTAC 4974
 Db 5002 TTCACTCACGGAAGTAC 5018
 RESULT 10
 VEU5347
 LOCUS
 DEFINITION Venezuelan equine encephalitis virus strain 6119 nonstructural polypeptide and structural polypeptide genes, complete cds.
 ACCESSION U5347
 VERSION U5347.2 GI:20800451
 KEYWORDS
 ORGANISM
 SOURCE
 ORGANISM
 Venezuelan equine encephalitis virus
 Venezuelan equine encephalitis virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus; VEEV complex.
 REFERENCE
 1 (bases 1 to 11423)
 Weaver,S.C., Salas,R., Rico-Hesse,R., Ludwig,G.V., Oberste,M.S., Bohehl,J., and Tesh,R.B.
 Re-emergence of epidemic Venezuelan equine encephalomyelitis in South America. VBE Study Group
 Lancet 348 (9025), 436-440 (1996)
 JOURNAL MEDLINE
 PUBMED
 8709783
 2 (bases 1 to 11423)
 Weaver,S.C.
 Direct Submission
 Submitted (18-APR-1996) Scott C. Weaver, Pathology, Univ. TX Med. Branch, 301 University Ave., Galveston, TX 77555-0609, USA
 3 (bases 1 to 11423)
 Weaver,S.C.
 Direct Submission
 Submitted (01-MAY-2001) Scott C. Weaver, Pathology, Univ. TX Med. Branch, 301 University Ave., Galveston, TX 77555-0609, USA
 COMMENT
 Sequence update by submitter
 On May 15, 2002 this sequence version replaced gi:2335117.
 FEATURES
 source
 1..11423
 /organism="Venezuelan equine encephalitis virus"
 /mol_type="genomic RNA"
 /strain="6119"
 /db_xref="taxon:11036"
 /note="antigenic IC"
 25..7506
 /note="contains low efficiency read through at stop codon; encodes nsp1, nsp2, nsp3 and nsp4 proteins"
 /codon_start=1
 /transl_except="(pos:5662..5664,aa:OTHER)
 /product="nonstructural polypeptide"
 /protein_id="AA028637.1"
 /db_xref="GI:20800451"
 /translation="MERVHVIDIEEDSFFLRALQSPQFEVAKQVTDNDHNAKAPS
 HLAASKLIEVDPSDDTIDIGSAPARMYSGRKHCICPKRCADBPRLKYATKLK
 NCKEITDEKDKKEELAAVMSDDLEETMCLHDDSCRYEGQVAVYQDVYAVDGP
 SLVQANKGVAVAVIGFDITPPFPIKULAGAYPEYSINWMDVETLTARNIGLGSDDV

[illegible]

Best Local Similarity 64.5%; Pred. No. 0;
Matches 3224; Conservative 0; Mismatches 1749; Indels 24; Gaps 3;

Oy	2	CCCTTCAAACTTAATGCAATCCAAATTGGAAAATTCACGTTGACCTTGAATGCTGACAGCC	61
Db	2	CCCAACCAATTAACCTACCCAAAAATGGAAAGTTCAAGTTGACATGAGGAAGACACCC	61
Oy	62	CGTATGTCAGTCCGTTACAGCGGACGTTTCCACAATTTGAGATCGAAGCAAGCAGCTCA	121
Db	62	CATTCCTCAGAGCTTTACACAGGAGCTTCCCGACGTTTGAGATGAGAACCAAGCAGGTCA	121
Oy	122	CTGACAAATGACCATGCGCAATGCGCAGAGCGTTTGGCATGTGGCAACAAAGCTCATTTAGA	181
Db	122	CTGATTAATGACCAATGCTAATGCCAGAGCGTTTGGCATGTGGCTTCAAAACTGATCGAAA	181
Oy	182	GCGAAGTCAGACCGGGAGCAAGTTATCTTTGGACATTTGGAAGGCGCCGCTCAGACATGCAC	241
Db	182	CGGAGGTGGAGCCATCCGACAGATCCTTGACATTTGGAAGGCGCGCCGCGACAGATGT	241
Oy	242	ATTCCAAATCAACCGGTATCATATTGATCTGCCCCATATGATPAAGCGCTGAGAACCCGACAGAC	301
Db	242	ATTCTTAAGCAATAAATATCATTTGATCTGTCCGATGAGATGTGGGGAAGATCCGACATATG	301
Oy	302	TACAAACGATATGCAAAAAGCTTAAAGAAAGT-----GACATTACGACAAAGACATAG	355
Db	302	TGTACAAGTATGCAACTPAAGCTGAAGAAAAATTCCAAGAAATPAATGACAAGGAATTTGG	361
Oy	356	CCTTAAGCGCGACACCTGCTGGAAGTCATGTCAACACGAGCGACAGACTCCATCTC	415
Db	362	ACAGAAAATPAAGAGAGTCCGCCCGCATAGACGACCCAGACTCGTGAACCTGAGACTA	421
Oy	416	TGTTATATGACACACAGCGCCAAGTATGATCTTTGGAAGTATGACAGTATACCAAGATG	475
Db	422	TGTGCTCTCAGAGAGATGAGTCAATGTGGCTPAAGAGGGCAAGTGGCTGTTACACAGATG	481
Oy	476	TGTACGAGTCCATGACACCGACATCAATCTACCAACGAGCGCTTAAAGAGTTAGACAA	535
Db	482	TATACGAGTTGACGCGACCGACAAGTCTCTATCACCAAGCCMAAGGGAATTTAGATCG	541
Oy	536	TTTATCGGATAGGCTTTGACACGACCCCTTATATGTAACAANAATGGCAGTTCCTAC	595
Db	542	CTTATCGATATAGGCTTTGACACACCCCTTATATGTTTAAACAACTTGGCTGAGCATATC	601
Oy	596	CTACTTACAAACGAACTGGGCTGACGAGAGATTTGGAAGCAAGTAACTTGGCCTCG	655
Db	602	CATATATCTACCAACTGGGCGACGAAACCGTGTTAACGGCTGTAACTAAGGCTTAT	661
Oy	656	GTTAACTCAGATCTTCAAGAGACGAGGCTTGGAAACTCTCAATCCTTATGGAAGAAAGGC	715
Db	662	GCACTCCGAGCATGAGAGCGGTACAGTAAAGGAGTGTCCATTTCTTGAAGAAAGTAT	721
Oy	716	TCCAACTCAATPAATAGATCAATTTTCGTTGGTTTCAAAATCTACACAAAGATAGAT	775
Db	722	TGAAACCAATCCAAATPAATGTCTATTTCTCTGTGGCTGACCATCTACACAGAAAGAGG	781
Oy	776	CACGTTAACGATGCGCATCTTCCAAACGTTTCCAATTGAAAAGAAAGTCTAATCTCA	835
Db	782	ACTTACTAGGAGAGCTGGCACCTGCGGTCTGATTTCACTTAAGTGGACAGCAAAATTA	841
Oy	836	CAGGTAGATGGACCAATTGTCACTGTGTAAGGGTATGTCATCAAAAAGATTAACATCA	895
Db	842	CATGTCCGGTGAAGACTATATGTTAGTTGCGACGGGTATCGTGTTPAAAAGAAATAGCTATCA	901
Oy	896	GCCCAAGACATAACGTTAAATTGAGAACTTGGGGTCCAAATCATGCGCAGGGGTTCT	955
Db	902	GTCACAGGCTGTATATGGGAAGCTTCAAGCTATATGCTGCTACGATGACCGCAGAGGATTC	961
Oy	956	TGAGTTGCAAGTACACAGATACGCTGCGCGCGACAGAGGTTCTTTTGTCTGTGTACGT	1015
Db	962	TGTGCTGCAAGTACACACACATTTGAACGGGAGAGGGTCTCTTTTCCGTGTGCAAGT	1022
Oy	1016	ATGTACCAAGCACCTTTCGATCAGATGACAGGAGATTTCTGGCAACTGACGTTATGTTGG	1075

Db 1032 ATGTCCAGCTACATTGTGTGAACAAATGACTGCACTGTGCAACAGATGTCAAGTCCGG 1081
Qy 1076 ATGAGCAACAAATCTATTGTGGGCTCAACCAAGATGTGCTGATGTGATGAGAGCG 1135
Db 1082 AGGAGCGGCAAAATCTGTGGTGTGGCTTAACGAGGCAATGCTGATCAACGGTCCGACCC 1141
Qy 1136 AAAGAAATCTAAACAATGCAAACTATCTATTACAGTGTGCGCCAGCGTTTCCA 1195
Db 1142 AAAAGAAACCAATACATGAAAGAAATATCTTTGGCCGTAGTGGCCAGGCAATTTGCTA 1201
Qy 1196 GGTGGGCGGTAACATCTGTCCGACTTGGACGACGAGAAAGACTAGGGGTTCGGAGAC 1255
Db 1202 GGTGGGCAAGAAATATAAGAAAGCAAGAAAGATGAGGGCCACTAGAGCTACAGATA 1261
Qy 1256 GCACTCTTACTATGGGCTGTGTGGGCTTTGCAAGCCCAAAATCAATCATCTCTCA 1315
Db 1262 GACAGTACTATGAGGCTGTGTGGGCTTTGAAAGGCAAGATTAATCTATTATTA 1321
Qy 1316 AGAAGCTGTGTACGCAAAATTAAGAAAGTACTGCGCTCTTTGACTCTATTGTGATTC 1375
Db 1322 AGCGCCAGATATCCAAACCATCATCAAAAGTGAACAGGATTTCCACTATTGCTGTGC 1381
Qy 1376 CAGCCCTTACCAACCAAGGCTGTGATGTGGCTTCCGCGTGAAGCTCAAGCTGTCTTG 1435
Db 1382 CCAGATAGGCAATACACACTGAGATGCGGCTGAAAGCAAGAAATCAAGAAATGCTAG 1441
Qy 1436 AACCACTGTCAACCCGACCGGCTATTACAATGGCCGATGTGAGCATCTGCGTGGT 1495
Db 1442 AAGACACAGAGGCGGCTCACTCTCATTAATGCGAGACATACAGAGGCTAAAGTGG 1501
Qy 1496 TACAGCAAGAACTGAAAGATGGCTGACGCGAAAGATCAGAGAGCCCTGCCACCT 1555
Db 1502 CAGCGGATGAGGCTAAAGAAAGTGGCTGAAGCCGAGAGAGTGGCGCTGACATCAACCTT 1561
Qy 1556 TGCTCCCTGAATAGAAAGAAAGACCGTGAAGGCAAGATGAGCTCATTAATGCAAGAG 1615
Db 1562 TGGCAGCTGATTTTGAAGAGCCACTGTGAAAGCCGATGTGCACTTGAATGTTTCAAGAG 1621
Qy 1616 CAGGAGAGGTAGGCTGAGACACCAAGAGACATCAAGGTGAACAATTTCCAGAGCG 1675
Db 1622 CTGGGCGCGCTCAGTGAAGACACCTCGGGCTTGAATTAAGGTATCAACAGTATGCGGG 1681
Qy 1676 AAGAGAAATTTGGCTTATACGCTATATCTTCAACCCAGCGGCTATTTGAATGAAAAAC 1735
Db 1682 AGGACAAATGGGCTTATGAGAGTCTTCTCCAGAGCTGTATCTCAAGATGAGAAAC 1741
Qy 1736 TGGCTGTATCCACCATTGGGGAACAAAGTACTGTAATGACTCAAAAGGTAGGCGAG 1795
Db 1742 TATCTTGCAATCACCTCTCGCTGAACAAGTCAATGATATACACACTGTGCGCAAG 1801
Qy 1796 GGAATACAAAGTCAAGCTATACACGCTAAGTCAATTTACAGAGAGGAGCGCGTCC 1855
Db 1802 GCGGTTATGCGCTGAACCTTACATGAAAGTGTGTGCAAGGAGCAATGCAATAC 1861
Qy 1856 CTGTTCAAGACTTCAGGCTATGAGTGAAGCGCTACGATCGTTTCAACGAGAGAGT 1915
Db 1862 CCGTCAAGAGCTTCAAGCTCTGAGTGAAGTGCACATCTGTGTAACAGAGAGT 1921
Qy 1916 TCGTAACAGATACCTGACACATCGCAATCAAGAGAGCGCTAAACACTGACGAG 1975
Db 1922 TCGTAACAGATACCTGACACATCGCAATCGAGAGAGCGCTGAACACAGATGAAG 1981
Qy 1976 AGTACTATTAAGCTGTAAGAGTCAAGACACAGACTCAGAAATACGCTTCAATTTGAGC 2035
Db 1982 AATATTACAAACCTGTCAAGCCAGAGACAGAGCGCAATACCTGTATGACATCGACA 2041
Qy 2036 CAGGAAAGTGTTAAGCGAGAAAGCGAGTCCCTTGTGCTAAACGGTGTGCTGAG 2095
Db 2042 GGAACCAATGCGCTCAAGAAAGATTAAGTACTGGGCTAGGCTTACAGCGAGCTGTGG 2101
Qy 2096 ATCCACATTTGACAGATTGGCTACAGAGTCTCAAGACAGACAGCAGCAGCTCACA 2155
Db 2102 ATCTCTCCCTTCATGAATTTGCTAGAGAGTCTGAGAACGTCGCGGCGCTCTTACC 2161

Qy 2156 AAGTCCCAACCATCGGAGTCTATGGAGTGCAGGTTCAAGTAAATCTGGAATCATCAAA 2215
Db 2162 AAGTACCAACCAATGAGGCTGTATGGCTGTGCGGGTCAAGCAAGTCTGGCATATTAAA 2221
Qy 2216 GCGCTGTGATTAAGAAAGATCTGTGTGAGTGCAGAAAGAAATCTGCGAATCA 2275
Db 2222 GCGGAGTCAACCAAAAGATCTGTGTGTGAGGCGCAAGAAAGAAATCTGCGAATTA 2281
Qy 2276 TCAAGGATGTAAAGAGATGAGAGTATGATGTGTGCTGTAGACTGTCCATTCAATGTC 2335
Db 2282 TAAAGGAGCTCAAGAAATGAAAGGCTGAGACGTCAATGCCAAGCTGTGACTAGTGC 2341
Qy 2336 TTCTAATGGGATTAAAGACCCCGTTAACCTGTGATCATTTGATGAGGATTTGCTGCG 2395
Db 2342 TCTTAATGATGCAAAACCCCGTGAAGACCTGTATATTTGACAGAGCTTTGCTGTTC 2401
Qy 2396 ATGAGGAGCGCTGTGCACTGATTTGCCATGTGCAACCTAAGAAAGTGTATGTGCG 2455
Db 2402 ATGAGGAGCTTCAAGAGCGCTCATGAGCATCATTAAGACCTTAAGAAAGGAGTGTCTGCG 2461
Qy 2456 GGGACCAAAACATGCGGCTTTCTTAACATGATGTGCTGAAGTACTTTAACCATG 2515
Db 2462 GGGATCCAAACAGTGTGCTTTTCAATATGATGTGCTGAAGTGCATTTTAACACAG 2521
Qy 2516 ACATATGACTGAAGTATACATAAAGCATCTAGAGAGGTGCACAGACTGTACCG 2575
Db 2522 AGATTGACAGGAGTCTTCAACAAAGCATCTGTGCTTGCACTTAATTCGTGACTT 2581
Qy 2576 CCATGCTTCCACGCTCTTCTTACGACAAAGCAATGAAGAGTAAACCATGTGCTGATA 2635
Db 2582 CGGTGCTCAACCTGTGTTTACGACAAAGAGATGAAGACGACGAAACCGAAAGAGACTA 2641
Qy 2636 AATATATATGATATCCACAGGACACAAAGCCGCAAAAGATATGATTTTAACT 2695
Db 2642 AGATTGAGATTGACCTCTGCACTGACATCAACCAAGCAGAGCATCTTCACTT 2701
Qy 2696 GTTTCAGAGATGGGTGAACAGCTACGATTGACTAACAATAACGAAATCATGACTG 2755
Db 2702 GTTTCAGAGGTGGTGAAGAGTTCGAAATGATTAACAAAGCAAGAAATATGACGG 2761
Qy 2756 CGGCTGATTCGCAAGGACTTACGCGGAAAGCGTTATGCTGTCAAGTAAAGTCAACG 2815
Db 2762 CAGCTGCTTCAAGGCGTGAACCGGTAAAGGCTGTATGCGGTGTGATCAAGAGTGAAG 2821
Qy 2816 AGAATCACTTACTGCGAGACTTGTGACACGTGAACGTGTACTTACACGACAGAA 2875
Db 2822 AAAATCCCTGTACCAACCACTCAGAACATGTGAACGTCTTCACTGACCCGACGAGG 2881
Qy 2876 AACGATGTCTGGAAGACGCTAGCTGTGATCCCTGATTAAGACACTTACAGTAAAT 2935
Db 2882 ACCGATGTGTGGAACCACTAGCGGTGATCCATGATTAAGAACTGAGAGGCAAGT 2941
Qy 2936 ATCCGCGGGAATTTACGCGCTTCAATTGACAGACTGCAACGCGCAACGACCATTAAG 2995
Db 2942 ATCTGGAATCTTCACTGTCCACGATAGAGATGGAAGCAAGAGATATGCACTATGA 3001
Qy 2996 CAGCGTCTTGATTAAGCCGACAGACAGTATGTCTCAAGATTAAGTGAACGTCTGCT 3055
Db 3002 GGCACATCTTGGAGAGACCGGACCTTACCGAGCTTTTCAAAATTAAGCGAAGCTGTGT 3061
Qy 3056 GGGCGAAGGCTTTGAGCACTGTGGCAGCGCAACATGTGCTGACGAGACGAGT 3115
Db 3062 GGGCCAAAGCTTTGTGCGGATGTAAGTGTGAGGCAATGTGACATGAAACAT 3121
Qy 3116 GGGAGACGCTTCAACCATTAAGCATGACAGAGCTACTCACTGAATAAGCACTGAAGT 3175
Db 3122 GGAACATGTGATTAATCTTCAAGACGAAAGCTCATCAGCAGAGATATGAAAC 3181
Qy 3176 TCTTTGACCAAGGTTCTTTGAGTGAACCTGAGACAGTGGGTATTTTCCGCTCTACCG 3235
Db 3182 AACTATGCGTGAAGGTTCTTTGAGCTGACAGCTGAGCTCGGCTATTTTCTCACACCACTG 3241

QY 3236 TCGCACTTACTTACAGGGATCAGACTGGGATTAATCTCGCAGGGAAGAACTATATGGGC 3295
 Db 3242 TTCCGTTATCCATTAGGAATATATCATGGAATTAATTCCTCCCTCCCTTACATATGACGGGT 3301
 QY 3296 TTAAATAGAGAGTACAAAGAGTGTCTACCGGCATATCCGTGATCACAAAAGCGGTG 3355
 Db 3302 TGAATTAAGAGTGTCCGACGCTCCGACAGTACCCCAACTGCTGCAGACAGTTG 3361
 QY 3356 ACACAGGAGGAGTACGATTAAGAAATTAATACATCAAGAGTACTCTCTCAACATTA 3415
 Db 3362 CCACCGAAGAGTCTATGACATGAATACTGGCAGCTGCGCAATTAATGATCCGCGATTA 3421
 QY 3416 ATGGTGTTCATTAATCGCCGTTGCCCACTGGTATGATGGTGAACCAAGAGACAG 3475
 Db 3422 ATCTAGTACTGTGAACAAAGACTGCTCTCATGCTTTAGTCTCTCAACATATGAACCC 3481
 QY 3476 GTACAACTGATCAGACGGGATTCCTATCTAATGAAGAGCAAACTGTGTGTGTGATCG 3535
 Db 3482 CACAGAGTACTTTCTTCAATTCGTGAGCAAACTGAAGGGCAGAACTGTGTGTGTG 3541
 QY 3536 GCGATCTATGACATTCAGAGGAAGAAAGTACGTCCATGGGTCCATGCTCCCACTAATA 3595
 Db 3542 GGGGAAGTGTCCGTCCAGCAAAAGGTGCACTGTGTGTGACACCAAGCTTGAAGCTTA 3601
 QY 3596 CCATCAGGTGTGATCTGATTTGGGAATACCTAGCATGTGCGTAATATGACATTAATCT 3655
 Db 3602 CTTTAGAGCTCGCTGATTTAGTATCCAGGTGACGTCCCAATATGACATTTGAT 3661
 QY 3656 TTGTCAATGTTAGAACCCCGTACAGAAACAATCACTACCAAGTGCAGAGATCAAGCTTA 3715
 Db 3662 TTATTAAGTGAAGACTCATATTAATATCATCAATTAATACAGTGTGAAGACACAGCCA 3721
 QY 3716 TCCACCAAGATCTAAGCTGTAAAGTGTCTACACCACTGAACACTGGCGGAACATGTG 3775
 Db 3722 TTAAAGCTTAAATGTGACCAAGAAAGCTGTGTGCAATTTGAATCCCGCGGAACCTGCG 3781
 QY 3776 TGGCTATAGGGTATGGGCTTCTGATCGCGCAACGAGAAATATATCATCTCGGTGAC 3835
 Db 3782 TCAGATATAGGTTATGTTAGTCTACAGGCGCAGCGAGACATCATTTGTCTATAGCGC 3841
 QY 3836 GCTCATTTAGGTTAACCGGTGTCTGTACAGCTTAAGAAACATGCGGAATATCTAGGTTG 3895
 Db 3842 GGCAGTTCAAGTTCTCCGGGTATGCAAAACGAATCTCTCATGAAGAGACAGAAAGTAC 3901
 QY 3896 TCTTCGTGTTCTTGGCAGAGACAAAGGCAACCAACATGACAGAGACAGACTCGGTG 3955
 Db 3902 TGTGTGATTCATTTGGTATGATCGATCGCAAGGCCCGTACCAATCTTAAAGCTTTTCA 3961
 QY 3956 TAGTGTGACATCATCATCAAGGGTCAACCAAGGTACGAGGACAGGAGAGCTCCAGGCT 4015
 Db 3962 CTACCTTGACCAATCTATACAGTTTCCAGATTCACGAAAGCCGATGCCACCTCTAT 4021
 QY 4016 ACAGAGTATCAGAGGTGACATTAAGCAAGCGCTGACCAAGCTATCTGTTATGCTGCTA 4075
 Db 4022 ATCATGTGTGTCGAGGGGATATTGCAAGGCCACGAAAGAGTATCATTAATCTGCGCA 4081
 QY 4076 ATACAAAGGTCACCAAGTTCCGAGTGTGCGGTGCACTGTACCGAAATAGCCGCGGTG 4135
 Db 4082 ACACCAAGGCAACCTGTGCGAAGGGGTGCGGAGCGCTGTAAAGAAATTTCCGGA 4141
 QY 4136 CTTTGTAGACAGCAATAGCTGTGCGGACGGCTAGACTGTGTAAGCAACCAAC--GC 4192
 Db 4142 GCTTCGATTTTACAGCCATGCAAGTGGAAAAAGCGCATGTCAAGGTGTAATCAAC 4201
 QY 4193 TCATCATACATGCTGTAGAACCAATTTTCTAAGATGCCGAACCGAGGGCGAAGCTTA 4252
 Db 4202 ATATTATTCATGCGGTAGGGCCCAACTTCAACAAAGTTCCGAAGTTGAAGGGACAAAC 4261
 QY 4253 AGCTGCGAGCTGCTACATGACATAGGCTATGCTTAAGCGCTGAGGGGATTACAAAA 4312
 Db 4262 AGTTGGCAGAGCTTATGATGATCGTAAATTTGTCMAAGATTAACATTTACAGTACG 4321
 QY 4313 TATAGTACCGCTACTGTCAACCGGCATCTATTCTGTGGCAAGATGAGTGTGAT 4372

Db 4322 TAGGATTTCACTGTTGTCCACCGGATCTTTTCGCGAAACAAAGATGACTAACCCAT 4381
 QY 4373 CATTGCATCACCTGTTTACTGCTTTTCAACACTACGATGCCGATGTACCATATATGCT 4432
 Db 4382 CATTGAACCAATTTGCTGACGCTTTAGACACCACTGATGAGATGTAGCTTATACGCA 4441
 QY 4433 TGAATTAACATGGGAGACCAAGATTAATGAGGCAATTCACCCGAAGAAAGCGTCGAA 4492
 Db 4442 GGGCAAGAAATGGGAATATCTCTCAAGAGACATGCTTAGAGAGAAAGCACTGAGG 4501
 QY 4493 TTCTGGA-----TGATGACAAAGCAATGACATTTGATGCTGACAGGCTCC 4537
 Db 4502 AGATATCATATACAGACACTCTTCGCTGACAGAACCGAGATGACAGACTGTGAGGGTAC 4561
 QY 4538 ACCCAACAGCTCTTTGGCAGGACAGACCAAGTATCTCCATATGAGGCAATGTAT 4597
 Db 4562 ATCCGAAGATTTCTTGGCTGGAAGGAGGCTTACAGCAAGTATGAGCAAGCTTTCT 4621
 QY 4598 CATACCTGGAAGGTACAGATTCATGACACCGCAAGACATTTGCCGAATTCATGCA 4657
 Db 4622 CATATTTGGAAGGACCAAAATTTCAACGCGCGCCAGATATAGCAAAATTAATGCCA 4681
 QY 4658 TGTGCCCCCAAAATCTGAGGCTTAATGACAGATTTGCTGTACATCTGGGGAGAGTA 4717
 Db 4682 TGTGCGCAAGTTCAGACGAGGCGCAATGACAGATATGATATCTCGGTGAAGCA 4741
 QY 4718 TGTCCAGATCCGCTCCAAATGCCAGTAGAGAGTACAGGCGTCTGCTCACCTACA 4777
 Db 4742 TGAGCAGCATTTAGTCCAAATGCTCCGTCGAGAGTGTGGAAGCTCCACACCATGCA 4801
 QY 4778 CACTTCATGCTGTGTATTAAGCTATGACGCGTGAAGCGCGTATACAGGTGCGCTCTG 4837
 Db 4802 CGCTGCTTGTGTGTGATCTCATGCTATGACTCCAGAAAGATACAGCCTAAAGCT 4861
 QY 4838 CGAAGAAAGAAAGTGTCCGCTATGCTCATCATTTCTGTTCCGGAATACAGATCAAG 4897
 Db 4862 CAGTCCAGAAACAAATTAAGTGTGCTCATCTTTTCAATTTCCGGAATATGAATCACTG 4921
 QY 4898 GCGTGCAGAAAGCTACAGTGCAGCAACCAAGTCTGTTTACAGGCGTGTACACCGGCTG 4957
 Db 4922 GTGTGCAAGAAATCCAGTGTCTCCAGCTTAATCTGTTCTACCGAAGGTGCTCGTACA 4981
 QY 4958 TACACCCAGAAAGTAC 4974
 Db 4982 TTCAATCCAGAAAGTAC 4998

RESULT 11
 EVNSPERA 11447 bp ss-RNA linear VRL 25-MAY-1999
 LOCUS DEFINITION Venezuelan equine encephalitis virus, complete genome.
 ACCESSION L01442
 VERSION L01442.2 GI:4887231
 KEYWORDS
 SOURCE
 ORGANISM
 Venezuelan equine encephalitis virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 Alphavirus; VEEV complex.
 REFERENCE
 1 (bases 1 to 11447)
 Kinney,R.M., Johnson,B.J., Brown,V.L. and Trent,D.W.
 Nucleotide sequence of the 26 S mRNA of the virulent Trinidad
 donkey strain of Venezuelan equine encephalitis virus and deduced
 sequence of the encoded structural proteins
 Virology 152 (2), 400-413 (1986)
 JOURNAL MEDLINE
 86263392
 3088830
 REFERENCE
 2 (bases 1 to 11447)
 Johnson,B.J., Kinney,R.M., Kost,C.L. and Trent,D.W.
 Molecular determinants of alphavirus neurovirulence: nucleotide and
 deduced protein sequence changes during attenuation of Venezuelan
 equine encephalitis virus
 J. Gen. Virol. 67 (Pt 9), 1951-1960 (1986)
 JOURNAL

[illegible]

mat_peptide 45. .1649
/gene="NS"
/product="nonstructural protein nap1"
/function="minus strand RNA synthesis; methyltransferase;
guanyltransferase"
1650. .4031
/gene="NS"
/product="nonstructural protein nap2"
/function="RNA helicase; nonstructural proteinase;
necessary for subgenomic 26S mRNA synthesis"
4032. .5702
/gene="NS"
/product="nonstructural protein nap3"
5703. .7523
/gene="NS"
/product="nonstructural protein nap4"
/function="RNA polymerase"
45. .5684
/gene="NS"
/codon_start=1
/product="putative nonstructural polypeptide precursor"
/protein_id="AAC19323.2"
/db_xref="GI:48872313"
/translation="MEKVYHVDIEDSPFLDALQSPFQFEVAKQYTDNDHNAARFSLHSLATLEIDVDPDITLUDISAPARMYRSIKTHKICCPKACEDVPRLLKATLKKNCKEITLDEIDPKKKEELAAVMSDDELTEMTKLBDESCYCEQVADVPYAVAVDPTSLHYQANGVAAVWVIGFDTPPEMFKNLCAAYSPYSNMADEIVLTPARNGICSDVMERSRGSILAEKTKLKSNNVLFSVGSTIYAEKRDLSRHLSPVFLRKKQNYTCRCETIYSCDQYVLRITAIISPLGXKPSGYAAVHREBPGLCCKVTDTLNGERSFPVCTVYPATLCDDMTGVLTVASADDAOKLVGLNRIYVNGEPTONTMTNKKYLLPVAQAPARMAKXEYEDDERPFLRLDRQLVMGCCAPRRHKITSITYKPDOTIITIKNSDHPFVLPRTGSATTEITGLRTIRKLEBHSKPSPLTAEVDQKSCADAEKEVERAEELSAALPLPLADVEEPLTAEVDLMLQDAGAGSVETPRGLIKTYSVAGEDKISYALBSQAVLKEBKSLSCHPLAEQVIVITGHSGRGRVAVEPHGVKVVPEGHAI PVDFOLSESA TTVYNERFVNRVYLHIATGAGLNTDEEYKTVKPEBHEGLVLDIDRKCQYKEELV TGLDGTGLVDPPEHFAVPSALPAPVOPVLIGYGVGSGKGIITKSAVTXKDLVSGAKKCEALITDYKMKMGKLDNATVYNSVLNKGKGLVETLITIDEAPACAGTLRALIATIRKKAVALCGDPKQCGFPMNMLKXHPNHEICTQVFKHSISRCTKSTSVS TLFDKRRATTPNPKETXIVITDGTSGTKPKODDILTCFERGVQLOLDYKGNIMTPA ASQGLRGRGVAAVRYKXENPLVAPTEHKNVLLTREDIWKTLADGEMWITLTAKEQNPTFTIEMQOEHAIMRHLIERPDPVQNKANVMCAALVPVLTAGIDMTT EONATVYFETEDKASAEIYVNLQCVFPGLDLSCFSAPVPLIRNNHNSPSP MNMGYKENVQSLRSPOLPRAVATRVDMNTGLRNDPTLNVPRRLPHALV LHHNHEQSDSSPFSKLTGRVLYVGEKTSVGVKMDLSDRBEATFVRRLDGLBG DVPYCDITL PNVNRTPYKHHYOCEBDPAITLSMLTKKACHLNPGSTCYGAYDR ASESTIGCAIARQFNRVCKPKSLEBTEVLVFTIGDRKARHPNPKUSTLNTIYTGSRILHEGAKYFHVHVGDIATATEGYVIAANASKQPGAGVGCAIYKFPESFDIOP LSGRIHEGAKYFHVHVGDIATATEGYVIAANASKQPGAGVGCAIYKFPESFDIOP LLEVKARLYMKAAGHIIHAVGPNPNKYSVEBGDQKAAEYSIAKLVINDVNVASVPI LSTGTFSGNDRLTLQSLNHLITLADTADVAIYCDKMEAMTILKEAVARAAREVAI CIGDSSVTEPDEALVHVHSPSSLAGKSGTSDGKTFVLEGTCKHMAQDLAEINA CMPPATSENBOVCVNLIGESMSIIBSKCPVSEASAPSTETLCTIHOAMPVORL KASRPEQUTYNSFPLPKRYITGVOKTQCSQPLIPKPAVYVHPKYLVEPNDET PEPSANOSTEGTGPPEPPLITBEDRTRTTEPILITEEBEDSISLSDGPTHQVLOVE ADIHGPPSVSSSSWISPHASDPVDQSLISLDTLEGASVTSAGTSALTSNTPFASKEFL RAPVAPRTYVNRPNHPAPRTTSLAPSAACPSLTVSPNNVNIITRELEALTP SRTSPRSYRSRLSNPNPGRVNRVITREPEFAFAAOO"

MPVATVETAAOCWVITIGESMSISKCSQSEPALESTPSTLPCICLHMTPEBVOBT
KASBPQIUTVCSFPLPKRTIRVQKQICQSSQPLSPKPAITVHPKQIVLTPVQVET
PEBPASNOSTEGTEPPPELITDETRTRTRPEPIIIEEEDSISLSDYKTHQVLOVE
ADIHGPPSVSSSSWSVPHASDPDVSLSLDTLDEGASVSGASATSNVSPKASMEFL
AAVPAFVFPVSNPHAPARTDPVSLPAPASGRTSLVSTPGVNAVITREELALTP
SHTPSSVSTSLVSNPQVNVITREEEPAFAOQORPDGAVITSSQGGHIOQ
KSVQVULSEVLRELEITVAPRLDQKEBELRKQKLOLPTTPNARKSOSKQVEM
KAITARILQOGHUTIKAGKVECTHLPVPLVSSVNAESPVAEACAMKLE
NPPVAOSYCIIEPYADVLMDVQASCDLDTASFCPKPLSPKQSVLETPISAVPS
AIQNTLQALITATKNCVQNTRELPVLVDASAFNVECPKVCNNMEYETPEANR
LREENVNITVILKQKAALAKHKTNRNLMLQIIPDRFVMDLRVQKTPGTHKEP
RKQVVOIADAPLTPALVTCGHELVTRNLVAILPHTLPMQASDPALVIAEHOP
GCVULETDLASDPKSEDDMALTAMLLEDYDAELVILIEBAFESISILPTK
FFGAMKSMGKMTLIDVNVNAPINVASRLRELSYSCAALFGDNIKGVASDKLM
ARQCATVMEKXILDAVVEKAPVCGGFIICDSYTCGACVADPLKRLFKGPKPA
ADEHDDRRRLALHEESTTWMNVGILSELCKAVESYETVGTIIIVAMTTLASVYS
FSLYCAPIPLTVG"
45. 1649
/gene="NS"
/product="nonstructural protein nsp1"
/function="mRNA strand RNA synthesis; methyltransferase"
/quantity="transferrase"
1650. 4031
/gene="NS"
/product="nonstructural protein nsp2"
/function="RNA helicase; nonstructural proteinase;
necessary for subgenomic 26S mRNA synthesis"
4032. 5702
/gene="NS"
/product="nonstructural protein nsp3"
5703. 5723
/gene="NS"
/product="nonstructural protein nsp4"
/function="RNA polymerase"
45. 5684
/gene="NS"
/codon_start=1
/product="putative nonstructural polypeptide precursor"
/protein_id="AAC19323.2"
/db_xref="GI:4887223"
/translation="MEKRVNDLEDSPLRLAORSPOEVEVAKQVTDNDHNAFAS
HLASKIETREVPDSPTIILDIGAPARMYSKKHICTCMRCAEDPRLIKATKLAK
NCKETLDELKAKMELAAVMSDPDLLETWCLHDESQREQVAVYDDVADVADPT
SLHQAOKNDKLVLAAPVTEFPMSKNLGAVPSTWNADEPVLTAARVILCSDSV
ESRSGMSILRKRYLKPNSNVLFPSGTYIHEKDLRSMHLPSPHLKQNYTCRC
ETTVSGDGVVRIASIPELYKSPSYATMBEGLCKKDTLNGEBSPPVCYV
PRTLDQNTGLATVSDADQKLVGLNQRIVANGRTORNTMKNTLYLPVAQAP
PRAKYEQDEBERPGLRDLWGCWCAFRKHTVTSYKRPDTOTLIKVSDFSP
VLPRISNTLELGLTRLRKMLEHKEPSPILTAEDVQAKAADBAKREAREBELRA
ALPLAADVEHLEADVDLMQEAQASVETPRGILKTVSAEGKISGVALSPOA
VLPKSELSCIEHLEADVDIYTHSGRGAFAVEPYGLKIVAPGEGHIVQDFQALSSA
ITLVNREENVNVLIHGHGALLNDEVTYVSEHEDGQILDKQCKEKL
TGLGTELVDPPEFAYEASLRTPRADYQVPTTGVYPSGSKGITIKSANTKDL
VVSAKENCAEILRYKPKKGLDVNARVYVSLVNGCKPVELVYIDEAFACHAGTLR
ALAIIRPKKAPLGGDPKQCGFNMCLVHFNHEICTOVFNKHSISRCKTSVSVS
FLTEPYKAKTTPNKETKQIVITDTGCTKQODDILICFGVAKLOIDGPMITAA
ASQGTARKGVNAVKNVNPVYAPRSEHNVAILRTEDRIKWLTAQPMITAK
VGNSTATITEEQDAHDAMHRLIRRPDTPVFOKNAQCMKALVLPGLKIGIMDT
EQMNTVDFETDKAHSAELVNLQCRFFGLDLSLSAFTPLSIRNNHNSPSP
NNYGLNKEVNEPSSRYPOLPAAVNAVYDNNKTILNYPDILNVPNRRLPAPLV
LHNNHNPQSPFSPKSGRVLVNGEKLSPGKNVIMLSPKATPPARLDLGI
DVPKADILFVNARTPKVHYQOCEDEHAKLSMLTKKACLHNPQGTCSVIGYVAD
ASESILGAIRVAFKRSRYCKPSLEETEVLVPTIGDKKARHNHYKLSLTINLYT
GSRLEHAGCAEPHYVHGLIARATGCVIINAGSKQOQCGVQALYKRPSPSFDLP
IEVGRALYKCAKAITHAAGVNPKNVSEYQKQOLAEYSAKITVNDNTKSAVL
LUSTIIFSNKQKRLTQSLNHLITJALDITDADVAICRDKMEWTLKEAFAAREVAEII
CISDVSVEYPAELVYRVAHSGKSGKSYSDGTFSEYETKTHQAKQIAEINAA
MNPVATEANQCVIILGSMSSISKCSQSEVASESTPSTLPCICLHMTPEBVOBL
KASBPQIUTVCSFPLPKRTIRVQKQICQSSQPLSPKPAITVHPKQIVLTPVQVET
PEBPASNOSTEGTEPPPELITDETRTRTRPEPIIIEEEDSISLSDYKTHQVLOVE
ADIHGPPSVSSSSWSVPHASDPDVSLSLDTLDEGASVSGASATSNVSPKASMEFL
AAVPAFVFPVSNPHAPARTDPVSLPAPASGRTSLVSTPGVNAVITREELALTP
SHTPSSVSTSLVSNPQVNVITREEEPAFAOQORPDGAVITSSQGGHIOQ
KSVQVULSEVLRELEITVAPRLDQKEBELRKQKLOLPTTPNARKSOSKQVEM
KAITARILQOGHUTIKAGKVECTHLPVPLVSSVNAESPVAEACAMKLE
NPPVAOSYCIIEPYADVLMDVQASCDLDTASFCPKPLSPKQSVLETPISAVPS
AIQNTLQALITATKNCVQNTRELPVLVDASAFNVECPKVCNNMEYETPEANR
LREENVNITVILKQKAALAKHKTNRNLMLQIIPDRFVMDLRVQKTPGTHKEP
RKQVVOIADAPLTPALVTCGHELVTRNLVAILPHTLPMQASDPALVIAEHOP
GCVULETDLASDPKSEDDMALTAMLLEDYDAELVILIEBAFESISILPTK
FFGAMKSMGKMTLIDVNVNAPINVASRLRELSYSCAALFGDNIKGVASDKLM
ARQCATVMEKXILDAVVEKAPVCGGFIICDSYTCGACVADPLKRLFKGPKPA
ADEHDDRRRLALHEESTTWMNVGILSELCKAVESYETVGTIIIVAMTTLASVYS
FSLYCAPIPLTVG"

```
old_sequence 106..107
/gene="NS"
/note="this correction changes amino acid nsP1-21-Thr to
nsP1-21-Ser, this 1992 data corrects the 1989 reference in
citation [3]"
/citation=[3]
/replace="gc"
1533..1534
/gene="NS"
/note="this correction results in changing amino acid
nsP1-497-Arg to nsP1-497-Ala, this 1992 data corrects the
sequence reported in the 1989 reference in citation [3]"
/citation=[3]
/replace="gc"
1696
old_sequence
/gene="NS"
/note="this correction results in changing amino acid
nsP2-16-Asp to nsP2-16-Ala, this correction adds another
difference between the Trinidad donkey and vaccine TC-83
strains reported in citation [3]"
/citation=[3]
/replace="c"
17532..11444
/gene="26S mRNA"
7532..11444
/mRNA
/gene="26S mRNA"
/note="the structural proteins of the virus are translated
as a polypeptide precursor molecule from an intracellular
26S mRNA species that is identical to the 3'-one-third
portion of the viral genomic mRNA; putative"
/function="intracellular, subgenomic viral mRNA species"

Query Match 42.4%; Score 2119.6; DB 14; Length 11447;
Beet Local Similarity 64.6%; Pred. No. 0; Mismatches 1744; Indels 24; Gaps 3;
Matches 3220; Conservative 0;

10 ACTATGCATCCATATGGAAGAATTTCACGTTGACTAGATGTCAGACGCCGATATGC 69
30 AATTACCTACCCAAATGGAAGAGTTTACGTTGATCGAGAGAACGCCATTCCTC 89
70 AAGTCGTACAGCGGACGTTTCCAAATTTGAGATCGAAGCAGGACAGTCACTGACAT 129
90 AGAGCTTTGACGCGAGCTTCCGCAAGTTGAGTAGAAGCCAGCAGTCACTGATAT 149
130 GACCATGCCATGCGCAGGCGTTTTCGATGTGGCAAAAGCTCATTTGAGAGCAATC 189
150 GACCATGCTATGCGAGCGTTTTCGATCGCTGCTTCAAAATGATGAAACGAGGTG 209
190 GACCGGACCAAGTTATCTTGACATTTGAGAGTGGCGCCGTCAGACATGCACTTCAT 249
210 GACCATCCGACAGATCTTGAATTTGAAAGTGGCGCCCGCCGCAAGATGATTTCTAG 269
250 CACCGCTATCATTTGATCTGCGCTATGATTAAGCGCTGAAGACCGCGACAGACTCA 309
270 CACAAGTATCATTTGATCTGCTCGATGAGATGTGGCAAAATCCGACAGATGTAATAG 329
310 TATGCAAGAAAGCTTAAGAAAAGT-----GACATTTACGACAGAAATATGCGCTT 363
330 TATCACTAAGCTGTAAGAAAACTGTAAGAAAAATGATTAAGAAATTTGACAAAGAA 389
364 GCGGACAGCTGCTGGAAGTCAATGTCACACGAGCGAGAGACTCCATCTGTGTATG 423
390 ATGAAGAGCTGCGCGCGCTATGAGCGACCTGACCTGGAACCTGAACTATGTGCTC 449
424 CACACAGACGCAAGCTGTAGGTACTTTGAAAGTGAAGATTAACAGATGTGTATGCA 483
450 CACGACGAGAGGTGTGCTGCTACGAAAGGCGAAGTGTGTTTACCAAGATGTAACGG 509
484 GTCCATGACGCACTCAATCTACCAAGCGGCTTAAGAGATTAAGCAATTTACTGCG 543
510 GTTGACGACCGCAAGATCTATCAACCAAGCAATTAAGGAGTTAGTGTGCGCTACTGG 569
544 ATAGGCTTTGACAGACCCCTTTATGTAACAATAATGAGCGATTTCTACCTACTTAC 603
```

```
Db 570 ATAGGCTTTGACAGACCCCTTTTATGTTAAGAACTGTGCTGAGCATATCATCTAC 629
Qy 604 AACACGAACTGGGCTGACGAGAGATTTGAAAGCACTGAACTTGGCTTCGTAATCA 663
Db 630 TCTACCAACTGGGCGGCAAGAACGCTGTAAACGCTCTGTAACTAGGCCCTATCAGCTCT 689
Qy 664 GATCTTCAAGAGACAGGCTTTGAAAACTCTAATCCTTTAGAAAGAGAGGCTCAACT 723
Db 690 GACGTTATGAGAGGCTGACGTAAGAGATGTCATTTCTTGAAGAGATTTGAAACCA 749
Qy 724 ACTAATAGATCATATTTCTGCTGTTGTTCAACATCTACACAGAAATAGATCACTGTTA 783
Db 750 TCCACAAATGTTCTATCTCTGTTGGCTGACCATCTACACAGAAAGAGAGACTTACTG 809
Qy 784 CGTACTGCGATCTTCCAAACGTTTCCATCTTGAAAGAAAGTCACTTCAACAGTGA 843
Db 810 AGGAGCTGGACCTGCGCTGTATTTCACTTACGTGGCAAGCAAAATTAACATGTGCG 869
Qy 844 TGTGGACCATTTGACGTGTGAAGGCTGATCATCAAAAGATTAACATCAAGCCAGGA 903
Db 870 TGTGAGCTATAGTTAGTTAGTTGACAGGCTACGTTTAAAGAAATGCTATCATCTCCAGC 929
Qy 904 CTATACCGTAAAGTTGAGAACTTGGCGCTCCCAATGATCGAGAGGTTTCTTGAGTTGC 963
Db 930 CTGATGGAAGGCTTCAAGCTATGCTGCTACGATGACCGGAGGATTTCTGTGCTGC 989
Qy 964 AAGTCAAGATACGCTGCGCGGAGAGGTTTCTTCTGTGTATGATGTAACCA 1023
Db 990 AAGGTGCAACACATTTGAACGCGGAGAGGCTCTCTTTTCCGTGTGACAGTATGTGCCA 1049
Qy 1024 GCCACACTTGGCATGATGACAGGAGTTCTGGCAATCTGACGTTAGTGTGATGACGCA 1083
Db 1050 GCTATGATTTGTGACAAATGACTGCGATCTGGCAACAGATGTCAGTGGGAGAGCGC 1109
Qy 1084 CAATAATATTTGTTGGGCTCAACCAAGGATTTGCTCAATGATGTAAGACGCAAGAAAT 1143
Db 1110 CAATAATGCTGTTGGGCTCAACAGGATTTGCTCAACGCTGCGACCAAGAGAAAC 1169
Qy 1144 ACTAACCAATGCAAACTATCTATTAACAGTGTGCGCCGAGGCTTTTCAAGTGGCGC 1203
Db 1170 ACCAATACATGAATAATTAATCTTTGCGGTGTGCGCCGAGCATTTGTTAGTGGGCA 1229
Qy 1204 CGTGAACATGTCGCGCATTTGACGACAGAGAAAGAACTAGGAGGCGGAGCGCACTCT 1263
Db 1230 AAGGAATATTAAGAAAGTCAAGAAATGAAAGGCACTAGAGCTACAGATTAACAATTA 1289
Qy 1264 ACTATGGCTGCTGCTGGGCTTTCAGAACCCAGAAATCAATTCATCTACAAAGACCT 1323
Db 1290 GTCATGGGTTGTTGTTGGGCTTTTGAAGGACAAAGATTAATCTATTTTAAGCGCCG 1349
Qy 1324 GGTATGCAAAATTAAGAAAGTACCTGCTCTTATCATCTATTTGATTTCCAGCGCTT 1383
Db 1350 GATACCCAAACATCATCAAAAGTGAACAGGATTTCCACTCATTTGCTGCGCCAGATA 1409
Qy 1384 ACCAGCAGCGGCTGATATGAGGCTTCCGCGTATGCTCAAGCTGCTTGAACAACCT 1443
Db 1410 GGCAGTAAACATTTGAGATGCGGCTGAGAAACAGAAATCAGAAATGTTTAGAGAGACAC 1469
Qy 1444 GTCAAAACCCGACCGGCTATTTAATGAGCCGATGTGAGCATGCTGCGCTTACAGCAA 1503
Db 1470 AAGGAGCGGTACCTCTCATTAACGCGAGAGAGCTACAAAGATAGTGGGAGCGGAT 1529
Qy 1504 GAAGCTGAAGAGTGGCTGACGCGAGAGATCAAGAAAGCCCTGACCACTTCTCTCT 1563
Db 1530 GAGGCTAAGAGGCTGCTGAAGCGAGAGATTTGGCGCACTTACCACTTTTGGCAGCT 1589
Qy 1564 GAATTAAGAAAGAGACGTAAGAGGAGAGATGACCTCATTTATGCAAGAGGCGAGAGCA 1623
Db 1590 GATGTTAGAGAGCCCATCTTGAAGCGGATGTCACTTATGTTTACAAAGGCTGGGAGCC 1649
Qy 1624 GGTAGCGTGAAGACCAACGAGAGACATCAAGGTGCAAGTTTACCAGGCGAAGAGAG 1683
```

Dh 1650 GGCTCAGTGAGACACCTCGTGGCTTGAATAAGCTACGCTGCGAGACAAG 1709
Qy 1684 ATTGGGCTTACGCTATACCTTTCACCCGAGCGGATTGAATGAAAAATCGCGTGT 1743
Db 1710 ATCGGCTCTTACGCTGTGCTTCTCCGAGCGCTGATCTCAGAGTGAATAATCTTGC 1769
Qy 1744 ATCCACCATTTGGCGAAACAAGTACTGTATGACTCACMAAGTAAAGGCGAGAGATAC 1803
Db 1770 ATCCACCTCTCTGTAACAAGTCAATAGTATACACACTGCGCCGAAAGGGGCTTAT 1829
Qy 1804 AAAGTCAGCCATACCAAGGTAAGTCAATGTAACGAAAGGACGGCGCTCTTCA 1863
Db 1830 GCCGGAACCAATACATGTGTAAGTGTGCGAAGGACATGCAATACCCCTCAAG 1889
Qy 1864 GACCTTCAAGGCAATGAGTGAAGCGCTACGATGTTTTCAAGAGAGGAGTTGTAAC 1923
Db 1890 GACCTTCAAGCTCTGATGTAAGTGTCCACCATTTGTGTACAGAACGTGAGTTCGTAAC 1949
Qy 1924 AGATACCTGCAACCAATCGCAATCAACGAGAGCGCTAAACATGACGAAAGTACTAT 1983
Db 1950 AGGTACTGCAACCAATATTGCCACATGAGAGGCGGTGAACATGATGAATAATTAC 2009
Qy 1984 AAGATCTGAAGCTCAGAGACACAATCTCAAGATACGTTCTTGATTTGACGCAAGAA 2043
Db 2010 AAAATCTGAAGCCGAGGAGCACGCGCAATACCTGTACGACATGCAAGAAAG 2069
Qy 2044 TGTTGTAAGCGAAGACGAGTCCCTTGTGCTCAACCGGTGATCTGTGATGATCCACA 2103
Db 2070 TGCGTCAAGAAAGAACTAGTCACTGGGCTGAGGCTCACAGCGACCTGGTGTATCTCC 2129
Qy 2104 TTTCAAGATTGCTGACGAGAGTCTCAAGACAGACAGACGACCTCAAAATGCCA 2163
Db 2130 TTCCATGATTTGCTACGAGAGTCTGAGAACACAGACGCGCTCTTACCAATACCA 2189
Qy 2164 ACCATCGAGTCTATGAGTGCAGGTTCCAGTAAATCTGAAATCATCAAAAGCGCTGT 2223
Db 2190 ACCATGAGGCTGATAGGCGTGCAGAGATCAGCAAGTCTGCATATTAAGCCGCAATC 2249
Qy 2224 ACTAAGAAAGATCTGTTGTGAGTGCAGAAAGAAAGTGCAGAAATCATCGAGAT 2283
Db 2250 ACCAAAAAGATCTAGTGTGAGCCCAAGAAAGAAAGTGCAGAAATTAAGGAGAC 2309
Qy 2284 GTAAGAGATGAGACGATGATGATGCTGCTAGAGCTGCTGATTCAGTCTTAAAT 2343
Db 2310 GTCAAGAAATGAAGGCGTGAAGTCAATGCAAGATCTGAGATCTGATCTTGAAT 2369
Qy 2344 GGGGTTAAGCACCCCGTTAACTCTGTACATTTGAGGCAATTTGCTGCATGAGAG 2403
Db 2370 GGATGCAAAACACCCCGTAGAGACCTGTATATTGACGAAGCTTTGCTGTATGACAGT 2429
Qy 2404 ACGCTGCTGCACTGATGCTCAATCTCAAACTTAAGAAAGTGTATTTGTGGGGGACCA 2463
Db 2430 ACTTCAGAGGCTCATAGCCATTAATTAAGACCTAAAGGCGATCTCTGCGGGATCC 2489
Qy 2464 AAACATGCGGCTTCTTAAACATGATGCTGGAAGTACATTTTAACATGACATAGC 2523
Db 2490 AAACAGTGGCTTTTAAATGATGTGCTGAAGTGTCAATTTTAACAGAGATTTGC 2549
Qy 2524 ACTGAAGTACCATTAAGACATCTCTAGAGAGTGCACAGACGTAAACCGCATCTGC 2583
Db 2550 ACACAGTCTTCCACAAAGACATCTCTGCGGTTCACTAAATCTGTGCTTCCGTCGT 2609
Qy 2584 TCCAGGCTCTTCTAGACAGAGGATGAAGAGGTTAAACCATGTGCTGATTAATCTATC 2643
Db 2610 TCACCTGTTTAAAGCAAAATAATGAGAACGAGAAATCCGAAAGAGACTAAGATTTG 2669
Qy 2644 ATAGATCCACAGGAGCAAAAGCCGCAAAAGATGATCTGATCTTAACCTGTTTCA 2703
Db 2670 ATTGACCTACCGAGTACCAAACTTAAGCAGAGCGATCTTCTCACTTTGTTTCA 2729
Qy 2704 GGATGGGTGAACAGCTACAGATTGACTTCAAAATAATCAAGAAATCATGCTGCGCTCA 2763
Db 2730 GGGTGGGTGAAGCAATGATTAATTAAGCAAGAAATTAATGACGCGACAGCTCC 2789

Qy 2764 TCGAAGGACTTAAGCGGAAAGCGTTATGCTGACAGTACAAAGTCAAGAGATCCA 2823
Db 2790 TCTCAAGGCTGACCCGTAAGTGTGTATGCTGCTGCTGATGATGAATAATCT 2849
Qy 2824 CTCTACGCACTTCTGACACGTAAGCTGTTACTTACACGACGAAAAAGCAT 2883
Db 2850 CTGTACGCACTTCTGACACGTAAGCTGTTACTTACACGACGAAAAAGCAT 2909
Qy 2884 GTCTGGAAGAGCTATGCTGTGATTCCTGTGATTAAGAACTTACAGCTAAATATCCGG 2943
Db 2910 GTGTGGAAGAACTAGCGCGGACCCATGATTAAGAACTGATGCTCAAGATCCCTGG 2969
Qy 2944 GATTTACGCGCTTCTATGAGAGCTGAGCGGCAACGACGATTAATGAGCGCTT 3003
Db 2970 AATTTACTGCAAGATGAGAGATGAGACAGACATGATGCTATGAGGACATC 3029
Qy 3004 CTGTATGAGCGGACAGACAGCTGATGTGTCCAGAAATGATGAAGCTGTGCGGCAAG 3063
Db 3030 TTGAGAGACCGGACCTTACCGACGCTTCCAGAAATGAGAAAGTGTGTTGGCCAA 3089
Qy 3064 GCTTTAGGCGAGCTTGGCCACGCGCAACATTTGCTGACGAGACAGAGTGGAGAG 3123
Db 3090 GCTTTAGGCGGAGCTGAGAGACCGCTGCAATGACATGACCACTGAACTAATGAACT 3149
Qy 3124 TTGACCCATTCAGAGATGACAGAGCGTACCTGCAAAAGGCACTGATCTTTTGC 3183
Db 3150 GTGATTTATTTGAACGGAAGGCTCACTGAGAGATGATTAATCAACTATG 3209
Qy 3184 ACCAGTCTTTGAGATGAGCTGAGACAGTGGTATTTTCCGCTCTTACCGTGCAT 3243
Db 3210 GTGAGGCTCTTTGAGCTGATCTGAGCTCGCGTCTATTTTGTGACCACTGTTCCGTTA 3265
Qy 3244 ACTTACAGGATCAGACCTGAGATCTGCGCAGGAAAGAACTGATGAGCTTAAATGA 3303
Db 3270 TCCATTGAGAAATATCACTGAGATACCTCCGCTGCTCAATGATGAGGCTGAATAA 3329
Qy 3304 GAGGAGCAAGAGGTTGTACGCGGATATCGTGATCACAAAAGCGGTGACACAGG 3363
Db 3330 GAGTGTGCTGAGCTTCTGAGAGTACCAACACTGCTCGGAGTTGCCACTGA 3389
Qy 3364 AGGATGCTGATTAAGAAATATACATCAAGACTACTCCAACTAATATGAGT 3423
Db 3390 AGAGTCTATGACATGACATGTAAGTGTGCAATTAATGATCCGCGATTAACCTATGA 3449
Qy 3424 CCATTAATGCGCGGTTGCCACTGTTGATCTGTTGACCAAAAGCAGGGTACAACT 3483
Db 3450 CCGTAAACAGAAAGCTGCTCATGCTTATGCTTCCACATTAATGAACACCAAGAGT 3509
Qy 3484 GATCAGCGGATCTTATCTAAGATGAAGGCAAAATCTGTGTTGATGAGCGATCT 3543
Db 3510 GACTTTTCTTATCTGATGAGCAATTAAGGAGCAAGATGCTGTGTGTGGGAAAAAG 3569
Qy 3544 ATCAGATTTCCAGGAAAGAAAGTGAAGTCAATGGGTCAATGCGCACTAATACATCAG 3603
Db 3570 TTGTCCGCTCCAGGCAAAATGTTGATCTGTGTGCAAGCGGCTGAGGCTACCTTCA 3629
Qy 3604 TGTATCTGATTTGGAATACCTAGCAATGTGCTGTAATTAATGACATTAATCTTGTCAAT 3663
Db 3630 GCTGCGCTGATTTAGGATCCAGGTGATGCTGCTCAAAATGACATTAATTTGTTAAT 3689
Qy 3664 GTTAGAGCCCGTACAGAAACATATCATCAACAGTGCAGAGATCAGCTATCCACAC 3723
Db 3690 GTAGAGACCCATTAATTAATCACTATGAGAGTGTGAAGCAAGTGCATTAAGT 3749
Qy 3724 AGCATGCTAAGCTTAAGGCTGTCCACCACTGAACTGCGGAACTGTGTGCTTATA 3783
Db 3750 AGCATGTTGACCAAGAAAGCTGTCTGTGATGATCCGCGGAACTGTGTGACATTA 3809
Qy 3784 GGGTATGGCTTGTGATGCGCAACCGGAATATCATCTGCGGTGACAGCTCATTT 3843
Db 3810 GGTATGTTACGCTGACAGGCGGAGAAAGCATCTTGTGTGCTATAGCGCGCAGTTTC 3869

3844 AGTTTACCCGCTGCTGTGACGCTTAAGAACACTGCGCAAAATATGAGGTTCTCTGCTG 3903
3870 AAGTTTCCCGGATATGAAACCGAAATCTTCACTGAGAGAGAGAGTTCTGTTGTA 3929
3904 TTCTTCGGCAAGGACAAACGCAACCAACACATGACGAGACAGCTCGTGTAAGTCTT 3963
3930 TTCAATGGGTACGATCCGAAAGCCCGTACGCAACATCTTCAAGCTTTCATCAACCTTG 3989
3964 GACAACTCTTCAAGGCTCAACAGATACAGAGGAGGAGAGCTCCAGCCTACAGAGTG 4023
3990 ACCAACTTATTAAGGTTTCCAGACTCCAGAAAGCGGATGTGACCTCATATCATGTG 4049
4024 ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCGTTAATGCTGTAATAGCAA 4083
4050 GTGGAGGGGATATTTGCCACGCGCCACGAGAGATGATTAATGCTGCTACACGCAA 4109
4084 GGTCAACCAAGGTTCCGAGTGGCGGTGCACTGTACCCAAATGGCCGGCTGCTTTGAT 4143
4110 GGACAACTGGCGGAGGGGTGTGCGAGCGCTGTATAGAAATTCGCGAAAGCTTGAT 4169
4144 AGACAGCCAAATAGCTGTGCGGACGCTGACTGTGAAAGCAGAAC--GCTCATCTA 4200
4170 TTACAGCGGATCGAAGTGAAGAAAGCGGACTGTGCAAGGTGACAGCTTAAACATATTT 4229
4201 CATGCTGAGGACCAATTTTCTTAAGTGCAGGAAACCGAGGCGACCTTAAGCTGCA 4260
4230 CATGCCGATGACCAATTTTCAACAAAGTTTCGAGGTGTAAGGTGACAAACAGTTGCA 4289
4261 GCTGCTTACATGAGCATAGCTCATGCTCAACGCTGAGCGGATTAACAAATATCAGTA 4320
4290 GAGGCTTATGAGTCATGCTGATGATGTGTAACATTAACATTAACAGTACGATG 4349
4321 CCGCTACGTCAACCGGATCTATTTGTTGGCAAAAGTTCAGATGATCATTTGAT 4380
4350 CCACGTTGTCCACCGGATCTTTTCGCGAAACAAAGTGAATCAATCAATGTAAC 4409
4381 CACCTGTTCACTGCTTTCGACACTACGATGCGGATGTCACCATATTTGTTGGATTA 4440
4410 CATTTGCGACAGCTTTAGACACCACTGATGATGATGATGATGATGATGATGATGATG 4469
4441 CAATGGAGACCAAGATATGAGGCAATTCACCGCAAAAGGCTGAAATTTCTGGA- 4499
4470 AAATGGGAATGACTCTCAAGGAAGCAGTGTGAGAGAGAGAGTGAAGAGATATGC 4529
4500 -----TGATGACAGGCACTAGACATGCTGCTGTCAGGCTCCACCCAAAC 4545
4530 ATATCCAGCACTTTCAGTACAGAACTGATGAGAGCTGTGAGGCTCATTCGAAAG 4589
4546 AGCTCTTTGGAGGACAGCAGGTTACTCCGTCATAGAGGCAAGTTGATTTCAATCCTG 4605
4590 AGTTCTTTGGCTGAGAGAGAGGCTACAGCACAGGCAATGGCAAACTTTCTCATTTTG 4649
4606 GAAGGTACACGATTCATCAGACCGCCAAAGCAATGGCCAAATTCATGCAATGTGCCC 4665
4650 GAAGGACCAAGTTTCAACGAGCGCCCAAGATATGAGAAATTAATGATGATGAGGCC 4709
4666 AACAATATGAGGCTTAAGACAGATTTGCTTGTATCATCTCGGGGAGAGATGTCCAGC 4725
4710 GTTGCAAGGAGGCAATGACAGATGATGATGATGATGATGATGATGATGATGATGATG 4769
4726 ATCCGCTCAATATGAGGCTGAGAGAGTCAAGAGGCTGTGCTCCACTCACACATTTCA 4785
4770 ATTAGGTGAATATGCGCTGAGAGAGTTCGAGAGCTTCACACACCTTATGACGCTGCT 4829
4786 TGCCGTGTAAATTAAGCTATGACGCTGAGCGCTGATACAGGCTGCTGCTGCAAGAAA 4845
4830 TGCTGTGATCAATGACCATGCTCAGAAAGAGTACAGCGCTTAAAGCCTCACAGTCCA 4889
4846 GAACAGTTCCGCTTATGCTCATCTCTGTTGCGGAGTACAGATACAGGCGCTGAG 4905
4890 GAACAAATTAATGCTGTGCTCATCTTTTCATTTGCGAAGTATGATATCATGCTGTGAG 4949
4906 AAGCTACAGTGAAGCAACAGTCTGTTTTAGGCGTGTGATCAACCGGCTGTACACCCC 4965

Db 4950 AAGATCAATGCTCCCGCCTATATGTTCTCAACCGAAAGTGCCTGATATATCATCA 5009
Cy 4966 AGGAAGTA 4973
Db 5010 AGGAAGTA 5017
RESULT 12
AF004459 11420 bp RNA linear VRL 12-JUL-1999
LOCUS Venezuelan equine encephalitis virus strain 243937, complete
DEFINITION
ACCESSION AF004459
VERSION AF004459.2 GI:5442464
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
9261393
2 (bases 1 to 11420)
Wang, E., Barrera, R., Boshell, J., Ferro, C., Freier, J. E.,
Navarro, J. C., Salas, R., Vasquez, C. and Weaver, S. C.
Genetic and phenotypic changes accompanying the emergence of
epizootic subtype IC Venezuelan equine encephalitis viruses from an
enzootic subtype ID progenitor
J. Virol. 73 (5), 4266-4271 (1999)
JOURNAL
MEDLINE
PUBMED
10196323
3 (bases 8366 to 9182)
Powers, A. M., Oberste, M. S., Brault, A. C., Rico-Hesse, R.,
Schmura, S. W., Smith, J. F., Kang, W., Sweeney, W. P. and Weaver, S. C.
Direct Submission
Submitted (16-MAY-1997) Pathology, Univ. Texas Med. Branch,
Galveston, TX 77555-0609, USA
4 (bases 1 to 11420)
Wang, E. and Weaver, S. C.
Direct Submission
Submitted (12-JUL-1999) Pathology, Univ. Texas Med. Branch,
Galveston, TX 77555-0609, USA
Sequence update by submitter
On Jul 12, 1999 this sequence version replaced gi:3849804.
REMARK
COMMENT
FEATURES
source
location/Qualifiers
1..11420
/organism="Venezuelan equine encephalitis virus"
/mol_type="genomic RNA"
/strain="243937"
/db_xref="taxon:11036"
45..7502
/note="contains nonstructural proteins 1, 2, 3 and 4;
in-frame termination codon is read through at low
frequency"
/codon_start=1
/transl_except="(pos:5658..5660,aa:OTHER)
/product="nonstructural polypeptide"
/protein_id="AAD3357.1"
/db_xref="GI:5442464"
/translation="MEKVHYDIEDSEPFALORSFPOFEVAKQVTDNDHANAFAAS
HLASKLETVSDSDTLDIGSAPARMYKHXYHCISCPKACEDPRLKYATLKX
NKEITDKEIDKKKELAAVSDPDLETETMYKCHDSCBEGVAVYOVAVDPT
SLYHADKGVAVYVGFPTTPEMKTLAAVSYSTNMADDTYLTARNIGLSSDYM
ERSRKGSIIRKXIKYKRSNVVLSVSGTITHEKDLRSHVLSVFLRKQVYTC
ETTVSCGVYKAIASPGYKPSGVAATMHEBGLCKVCTDPLNGERSFVCTIV
PATLCDMTGILATDVADDAOKLLVGLNRIYVNGRTONTNTMKYLLPVAQAFA
RMAKEYEDQEDERPIGLRDROLVMCMAFRBHKITSYKRPDTYIILVNSDFHSF

[illegible]

BASE COUNT	3221 a	2821 c	2894 g	2484 t
ORIGIN				

Query Match	42.4%	Score 2119;	DB 14;	Length 11420;
Best Local Similarity	64.5%	Pred. No. 0;		
Matches 3220; Conservative	0;	Mismatches 1745;	Indels 24;	Gaps 3;

10 ACTAATCGATCATATGGAAGATTTCAGTTGACTTGGATGTCGACACCCGATATGTC 69

Db	30	AATTACTTACCCAAAATGAGAAAGTTACGTTGCATCGAGGAAGACGCCATTCCTC	89
Qy	70	AAGTCGTACGCGGACGCTTCCAAATTGAGATCGAAGCAAGGCACTGCACAT	129
Db	90	AGAGTTTGCACCGAGCTTCCCGAGTTTGAAGTGAAGGCCAAGCAGTCACTGATAT	149
Qy	130	GACCATGCCAATGCCAGAGCTTTTCCATGTGGCAAAAGCTCATTTAGAGCAAGTC	189
Db	150	GACCATGCTAACGCGCAGAGCTTTTCGATCTGGCTTCAAAATGATGAAACGAGGCTG	209
Qy	190	GACCGGGACCAAGTTATCTTGGACATTTGGAAGTGGCCCGCTCAGACATGCAATTC	249
Db	210	GACCATTCGACACGATCTTGACATTTGGAAGTGGCCCGCCGAGATGTTATTC	269
Qy	250	CACCGCTATCATTTGATCTGCGCTTATGATTAAGCGCTGAAGACCCGGAAGACTCAAC	309
Db	270	CATAGTATATCTTGATCTGTCCTGATGAGATGTGGGAAGATCCGGAACGATTTGTA	329
Qy	310	TATGCAAGAACTTAAGAAAAGT-----GACATTACCGACAAAGACATAGCCTCTA	363
Db	330	TATGCAATGATGTAAGAAAAAATTGCAAGAAAAATTACTGACAAAGAAATTGGA	389
Qy	364	GCGCGAGACCTGCTGGAAAGTATGTCAACACCAAGACGAGAGCTCCATCTGTGTATG	423
Db	390	ATGAAGGAGCTTGGCCGCGCTATAGCGACCTTGACTGGAAACGTAGACTATGTGCTTC	449
Qy	424	CACACAGACGCGCAGTGTAGTACTTTTGAAGTGTAGCAGTATACCAAGATGTATACGA	483
Db	450	CACGACGACGAGTATGTCTGCTTAGAAGGGCAAGTGCCTGTTTCCAGAGATGTATACGA	509
Qy	484	GTCATGCAACGGAATCAATCTTACCAACGAGCGCTTAAAGAGTTAGACAATTTACTGG	543
Db	510	GTTGACGGAACCGAACAAGTCTCTATACCAACCGACCAAGGAGTTAGAGTGGCTTACGG	569
Qy	544	ATAGGCTTTGACACGACCCCTTTTATGTATGACAAAAACATGCGAGTTCTTACCCTTA	603
Db	570	ATAGGCTTTGACACCCCTTTTATGTATGTTAAAGAACTTGGCGTGAATTCATCAATAC	629
Qy	604	AACACGAATCTGGGCGACGAGAGTATTTGAAAGCAGTAATTTGGCTCGGTAACTCA	663
Db	630	TCATCAACTGGGCGACGAAACCGTGTAAAGGCTGTAACATAGGCTTATGCAAGCTTC	689
Qy	664	GATCTTCAGAGAGCAGGCTTGGAAAACTCTCAATCCTTAGAAGAAAGGCTCCACCT	723
Db	690	GAGCTCATGAGCGGTCACGTAAGAGGATGTCCATTCTTBSAAGAAAGTATTTGAACCA	749
Qy	724	ACTAATAAGATCATATTTCTGGTTTCAACAACTTACACAGAGATAGATCACTGTTA	783
Db	750	TCCATATATGTCTATTTCTGTGGCTCGACCATCTTACACAGAAAGGAACTTACTTG	809
Qy	784	CGTAGCTGGCATCTTCCAAAGTGTGCCACTTGAAGAAAGTCTAATCTTACAGGTAGA	843
Db	810	AGGAGCTGGGACCTGCGCTGTATTTCACTTACGTGGCAAGAAAAATTAACAATGTGG	869
Qy	844	TGTGGGACCATTTGACGCTGGAAGGGTACGTCATCAAAAGATTAACATCAGCCCAAGA	903
Db	870	TGTAGACTATAGTTAAGTTAGTGGGACGGGTACTCGTTAAAGAAATAGCTATCACTCAG	929
Qy	904	CTATACGGTAAAGTTGAGAACTTGGCGTCCAAATGCAATGCGCAGAGGTTTCTTGAGTTGC	963
Db	930	CTGATATGGAAGCTTTCAGGCTATGTGCCACGATGACCCGAGAGGATTCCTTGATGTGC	989
Qy	964	AAATTCACAGTATAGCTTGGCGGGGAGAGGGTTCTTTTGTGTGTATGTAATGTACCA	1022
Db	990	AAATGTACAGCAATTTGAAGGGGAGAGGGTCTTTTCCCGTGTGACGTATGTGCA	1049
Qy	1024	GCCACACTTTGCGATCAGATGACAGGGATTTGCGCAACTGACGTTAGTGTGATGACGCA	1083
Db	1050	GCTACATTTGTGTACCAATATACGTGCACTGCGACACAGATGTCACTGTGGGAGACGCGC	1108
Qy	1084	CAAAATCATTTGTTGGGCTCAACCAAGATTTGCTCAATGGTAGAGCGCAAGAAAT	1143
Db	1110	CAAAATCTGCTGGTGGGCTCAACGAGGCTATCTGTCAACGGTGTGCAACCAAGAAAT	1168

QY 1144 ACTAACAAATGCGAGACTATCTATTACAGTGTGCGCCGAGCGCTTTTCAGGTGGCG 1203
DB 1170 ACCAAATACATGAGAAATTAATCTTTTCCCGTAGTGGCCGAGCATTTTCTAGTGGCA 1229
QY 1204 CGTGAACATGTCGCGCACTTGAGCAGCAGAAAGAACTAGGGGTGGCGAGCGCATCTT 1263
DB 1230 AAAGAAATATAGGAAGATGCAAGAGATGAGAGCCACTAGGAACTACAGATAGACATTA 1289
QY 1264 ACTATGGCTGTCTGTGGCTTTCAAGCCGAAATTCATCTCATCTCAAGAACTCT 1323
DB 1290 GTTATGGGGTGTCTGTGGCTTTTGAAGGCAAGATACATCTATTATTAAGCGCCA 1349
QY 1324 GGTACGAAACAATTAAGAAAGTACTGCGCTTTTGACTATTGTGATTCAGCGCTT 1383
DB 1350 GATACCAAAACATCACTCAAAAGTGAACAGCATTTCCACTATTGTGTGCCAGATATA 1409
QY 1384 ACCAGCCAGGGCTGATATGGGCTTCCCGTAGGCTCAAGCTGCTGTGAACCACT 1443
DB 1410 GGCAGTAAACATTTGAGATCGGGCTGAAACAAAGATTCAGGAAATGCTAGAAAGACGC 1469
QY 1444 GTCAACCCGCACTGGCTATTACATGGCCGATGTGAGCATCTGCTGGCTTACAGCA 1503
DB 1470 AAGGAGCGTCACTCTCATTTATGCGAGGACGTACAGAAAGCTAAAGTGGCAGCCGAT 1529
QY 1504 GAAGCTGAAGAGTGGCTGCGAGCGGAAAGATCAGAGAGCCGCGCACCTTCTCCCT 1563
DB 1530 GAGGCTAAGAGGTGCGTGAAGCCGAGAGATTTGCGCGCTGACTACCACTTTTGGCAGCA 1589
QY 1564 GAAATAGAAAAGAGACCGTAGAGGAGAAAGTAGACCTCATTTATGCAAGAGCAGAGCA 1623
DB 1590 GATGTTAGAGAGCCCATCTGTGAAGCCGATGTGATCTTTATGTCAAGAGGCTGGGGCC 1649
QY 1624 GGTAGCGTGAAGACCAACGAGGACATCTAGGTGCAAGTTACCAAGGAGAAAGAG 1683
DB 1650 GGCTCAGTGAAGACCTGTGGCTGTATGAAGGTTACAGCTATGCGGGGAGAGCAAG 1709
QY 1684 ATTGGGCTTTACGCTATCTTTACCCGAGCGGTATTTGAATAGTAAAGAACTGGCGTGT 1743
DB 1710 ATCGGCTCTTACGAGAGTCTTTCTTCCAGGCTGTACTCAAGATGAGAAATTAATCTTTC 1789
QY 1744 ATCCACCATTGGGGGAAACAAGTACTGTAAATGAATCAACAAAGGTAGGGCAGGAGATAC 1803
DB 1770 ATTCAACCTCTCGCTGAACAAGTCAATGATTAACAACCTCTGCCGAAAGGGCGTTAT 1829
QY 1804 AAGGTCAAGCATCTACAGGTAAAGTCAATTTACCAAGAGGAGCGGGGTCTCTTTCAA 1863
DB 1830 GCCGTGAACCTTACATGAAAAGTGTGTGCTCAGAGGAGCATGCAATACCCGTCCAG 1889
QY 1864 GACTTCAGGCAATGAGTGAAGAGCGCTACGATCGTTTCAACGAGAGGAGTTGTAAC 1923
DB 1890 GACTTTCAAGCTCTGAGTGAAGAGTGCACACTTGTGTCAACGAGAGTGTGTAAC 1949
QY 1924 AGATACCTGCACCAATCTGCATCAACGAGAGAGCGCTAAACACTGACGAAGTACTAT 1983
DB 1950 AGGTAATTGCAACAATTTGCAACATGAGAGAGCGCTGAACAAGATGAAGATATATAC 2009
QY 1984 AAGACTGTAAGACTCAGAGACACAGACTCAGAAATACGCTTCTGATTTGAGCAGCAAG 2043
DB 2010 AAAATTTTCAAGCCCAAGAGACACAGCGGGAATACCTGTATGACATCGACAGGAAACAG 2069
QY 2044 TGTGTTAAGCGAAGAGACGAGAGTCCCTGTGCTCAACCGGTGATCTGTGATTCACCA 2103
DB 2070 TGCCTCAAGAAAGAACTAAGTCACTGGGCTAGGGCTTACAGCGAGCTGTGTGATCTTCC 2129
QY 2104 TTTTACGAGTTTGGCTAGAGTCTCAAGACAGACGACGAGACCTTCAAAAGTCTCCA 2163
DB 2130 TTTCATGAATTCGCTTACGAGAGCTGAGAACAGACGCGCGCTTATCAAGTACCA 2189
QY 2164 ACCATCGAGTCTTATGAGTGCAGGTTCAAGTAAATTTGAAATCATAAAGCGCTGTG 2223
DB 2190 ACCATAGGGGTATATGGGTGCGGGTTCAAGCAAGTCTGCAATCAITTAAGGCGCAGTG 2249

QY 2224 ACTAAGAAAGATCTGGTGTGAGTGGCAAGAGAAATCTGCCAGAAATCATCAGGAT 2283
DB 2250 ACCAAAAAGATCTGGTGTGAGTGGCCAAAGAAAGAACTGGCAGAAATTAAGGAGC 2309
QY 2284 GTAAGAGATGAGAGTATGATGTTGTGCTGTAGAGCTGTGCATTCAGTCTTCTAAAT 2343
DB 2310 GTCAAGAAATGAAAGGCTGAGCGTCAATGCAAGACGTGTGACCTCAGTCTTGAAT 2369
QY 2344 GGGTTAAGCAGCCCGTTAACTCTGTACATTTGATGAGGCAATTTGCTGCCATGACAGG 2403
DB 2370 GATGCAAAACACCCCTAGAGACCTGTATATTGACGAAAGCTTTTGTCTGTACAGGC 2429
QY 2404 ACGCTGTGCACTGATTTGCAATGTCAACCTTAAGAAAGTGTATTTGGCGGAGCCA 2463
DB 2430 ACTCTGAGAGCGCTCATAGGCATCATTAAGCCTAAAGAGCAAGTGTCTGCGGGATCCA 2489
QY 2464 AAACATGCGGCTTTCTTAACTGATGTGCTGAAAGTACATTTTAACCTATGACATATGC 2523
DB 2490 AAACAGTGGCTTTTCAATATGATGTGCTGAAAGTGAATTTTAAACGAGATTTGC 2549
QY 2524 ACTAAGTATACATTAAGCATCTCTAGGAGGTGACACAGACTGTAAACCGCATGCTC 2583
DB 2550 ACGAGGCTTTCAACAAAGACATCTCTGTGCTGACTAAATCTGTGACTTGGTGTGTC 2609
QY 2584 TCCAGCTCTTTTACGACAAAGCAAGTAAGACGCTTAACCCATGTGCTATTAATCATC 2643
DB 2610 TCAACCTGTTTACGACAAAGAAAGATGAGAACAGCAACCGAAAGACTTAAGATTGAG 2669
QY 2644 ATAGATACACAGGAGACACAAAGCCGACAAAGATGATTTGATTTTACCTGTTTACA 2703
DB 2670 ATTGACACTATCTGCACTACTTAACAAAGCAGAGATCTCATCTTGTTCAGA 2729
QY 2704 GGAATGGTGAAGAGCTACAGATTTGACTACAAATATCAAGAAATCATGCTGGCTGCA 2763
DB 2730 GGGTGGTAAAGCATGTTGCAATAGATTACAAAGCAGCAAGAAATATAGCGGAGCTGCC 2789
QY 2764 TCCGAAGCACTTACGCGGAAAGCGTTTATGTCTGACGTACAAAGTCAACGAAATCCA 2823
DB 2790 TCTCAAGGCGTGAACCGTAAAGCGTGTATGCCGTTGGTATCAAGGTAAGTAAATCCC 2849
QY 2824 CTCTACTGCGACCTTGTGAGCACGTGAACGTGTACTTACAGCACAGAAAAACGATT 2883
DB 2850 CTGTACGCAACCACTCAAGAACATGTGAACGTCTACTGACCGGTACGAGGAGCCGATC 2909
QY 2884 GTCTGAAGAGCGTACGTGTGTATCCCTGATTAAGCACTTACAGTAAATATCCCGGG 2943
DB 2910 GTGTGAAAACATGACCGCGCATTCATGATTAAGAACTGACGCGCAATATCTCGGG 2969
QY 2944 GATTTCACGGCTTCATTTGAACGACTGCGAGCGGAAACGAGCCCATTTATGGCACGGGTT 3003
DB 2970 AATTTCATCTGCACGATAGAGATGAGCAAGAGATGAGGCAATCATGAGGACATC 3029
QY 3004 CTGTATAGCGCGAGACGCTGATGTGTTCAGAAATAGTGAACGTCTGTCTGGCGAAG 3063
DB 3030 TTGAGAGAACCGGACCTTACGACGTAATTCAGAAATAGGCGAAGCTGTGTGGCCAAAG 3089
QY 3064 GCTTTAGGCAAGCTTGGCGCACGCGCAACATTTGTGTCAGAGACGACGAGTGGAGAGC 3123
DB 3090 GCTTTGTGCGGTACTGAAAACCTGACGAGCATGACATGACCACTGAAACATGGAACACT 3149
QY 3124 TTGACCCATTCAAGACATGACAGAGCTACTACCTGAAATGGCACTGAACCTTCTTTTGC 3183
DB 3150 GTGGATTAATTTGAAAGAGCAAAAGCTCACTGACAGAGATGATGAAACCAACTATGTC 3209
QY 3184 ACCAGTTCTTTGAGATGACCTGTGACAGTGGGTATTTTCCGCTCTACCGTGCACCTT 3243
DB 3210 GTTAGGTTCTTTGAGACTGACCTGSACTCCGGCTATTTTCTGACCCACTGTTCCTGTA 3269
QY 3244 ACTTACGAGATCAGCATGGGATTAACCTGCGAGGAGAAAGATGTATGGGCTTAATAGA 3303
DB 3270 TCAATTGAGATATATCATGAGATATTCCTCGCTCAACATGATGAGGCTGAATTAAG 3329
QY 3304 GAGGTACAAAGAGTGTGACCGCGCATATCCGTGATCAAAAGCGTTGACACAGGC 3363

```

Db      3330  GAAAGGTCCTCCCTCAGCTCTCCCGCAGGTACCCACAGACTGCTCGAGCAGTGTCTACTGGA 3389
Qy      3364  AGGGATGCTGATATAGGAATAATACATCAAGACTACTCTCCAACTTAATGTGTT 3423
Db      3390  AGAGCTATATGATGATAACCCGCGCAGCTGTGCAATATGATCCCGCAATATCTAGTA 3449
Qy      3424  CCATTAAATCCGCGGTGCGCCACTGCTGTGATGCTGTGACCAAGAGACAGGGTCAACT 3483
Db      3450  CCAGTAACAGAAAGCTGCTCTATGCTTTAGTCTCTCCACCAATAGAACCCACAGACT 3509
Qy      3484  GATCAGAGGGATTCCTATCTAAGATGAAGGGCAAACTGTGTGTGTATCGGCACTCT 3543
Db      3510  GACTTTTCTTCACTGTGTCAGCAAACTGAAGGGCAGAACTGTGTGTGTGTGCGGAGAA 3569
Qy      3544  ATCAGCATTTCCAGGGAAGAAAGTATAGTCCATGGGTCCATTGCCCACTAATACCTACG 3603
Db      3570  TTGTCCGTCGCCAGGAAAACGTGTGATGTGTGACAGCACGCTGAGGCTACCTTCAGA 3629
Qy      3604  TGTGATCTGATTTGGGAATACCTAGCCATGTGCGTAAATATGACATTATCTTGTCAAT 3663
Db      3630  GCTCGCTGATGATTTAGTATCTCCAGTGAAGTCCCAATATGACATTTATTTAT 3689
Qy      3664  GTTAGGACCCGCTACAGGAACCATCACTACCAAGTGCAGATCACGCTATCCACAC 3723
Db      3690  GTGAGACTCCATATAAATATACCATCATTAACAGAGTGAAGACACGCCATCAAGCTT 3749
Qy      3724  AGCAGCTACAGCTGAAGGCTGTCCACCACTGAAACCTGGGGAACATGTGTGCTATA 3783
Db      3750  AGCATGTGACCAAGAAAGCTTGTGTGATGTGAATCCCGGGAACCTGCTCAGCAATA 3809
Qy      3784  GGGTATGGGCTTGTGATCGCGCAACCGAATATCATCTCGCGTGGCACTCATTT 3843
Db      3810  GGTATGTGTATGCTGACAGGGCCAGCAAGCATTTATGTGTCTATAGCGGCGAGTTC 3869
Qy      3844  AGGTTTACCCGCTGTCTGACGCTAAGAACATCTGCCGAAATATCTGAGTTCTCTTCG 3903
Db      3870  AAGTCTCCCGGGTATGCAACCGAATCTCTCATTTGAAGAGACAGAAAGTCTGTTTGA 3929
Qy      3904  TTTCTTGGCAGAGCAACGCGCAACACACATGACAGAGACTCGGTGTAGTCTT 3963
Db      3930  TTCAATTGGGTAGATCGAAGGCCCGGTAGCGCAAACTCTTACMACTTTCATCTACTTG 3989
Qy      3964  GACAACATCTATCAAGGGTCAACAGGTACGAGGAGGAGGAGCTCCAGCTACAGAGTG 4023
Db      3990  ACCAATCTATACAGGTTCAGACTCCAGAACCGGATGCGACCTCATATCATGTG 4049
Qy      4024  ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCGTTAATGCTGCTAATAGCAA 4083
Db      4050  GTGCGAGGGGACATTGCCACGGCCACCGAAGAGTATCATTAATGCTGTCTAACGACAAA 4109
Qy      4084  GGTCAACCAAGTTCGGAGTGTGCGGTGACCTGTACCGAAATGGCCGGCTGCTTTGAT 4143
Db      4110  GGACAACTGTGCGAGGGGTGTGCGAGCGCTGTATTAAGAAATTTCCGAAAGCTTCAT 4169
Qy      4144  AGACAGCAATAGCTGTGCGGAGCGCTAGACTGTGAAGACAGAAC---GCTCATCTA 4200
Db      4170  TTACAGCCGATCGAAGTGGAAAAAGCGACTGTGTCAAAAGTGTACAGTAAAGATATCT 4229
Qy      4201  CATGCTGATAGAACCAATTTTCTTAAGATGCGGAAACCGAGGGGCACTTAAGCTCCGA 4260
Db      4230  CATGCGTAAGGGCCCACTTCAACAAAGTTTCGGAAGTTGAAGTGAACAAACAGTTGGCA 4289
Qy      4261  GCTGCTTACATGAGATAGCTTCATGCTCAACGCTGAGCGGATTAACAATAATACGTA 4320
Db      4290  GAGGCTTATGAGTCACTGCTAAATTTGTCAACGATTAACATTAAGTGTAGTCCGATT 4349
Qy      4321  CCGCTACTGTCAACCGGATCTATCTGTGGCAAGATCGAGTACATCATTTGAT 4380
Db      4350  CCAGTGTGTTCACCGGATCTTCTTCGGAACAAAGATCGATTAACCAATCATTTGAC 4409
Qy      4381  CACCTGTCTACTGCTTTCGACATACGATGCGGATGTCAACCATATATTTGCTGATAAA 4440

```

```

Db      4410  CATTTGCTGACAGCTTTTAGACACCACTGATGAGATGTAGCCATATACTGACAGGACAAG 4469
Qy      4441  CAATGGGAGACAGGATTAATGAGGCCATTTCACCCCAAGAAAGGTTGMAATTTCTGAT 4500
Db      4470  AAATGGGAAATGACTCTCAAGGAGACAGTGTCTAGGAAGAACGATGAGAGATATGT 4529
Qy      4501  GATGACAAAGCC-----AGTAGCATTTGACTTGTGAGGGTCCACCCAAAC 4545
Db      4530  ATATCCGACGACTCTTCGGTACAGAACCCGATGACAGAGCTGTGATAGGATCATCCGAA 4589
Qy      4546  AGCTTTTGGAGGAGACCAAGTTATCTCCGTCAATGAGGCAAGTTGATTCATACCTG 4605
Db      4590  AGTTCTTTGGCTGGAAGAAAGGCTACAGCAACAGCGATGCGCAACATTTCTCATATTG 4649
Qy      4606  GAAGTATACGATTCATACACACCGCCAAAGACATGCGCAATTCATATGCAATGTGGCCC 4665
Db      4650  GAAGGAGCAAAATTTATCAGCGCGCCCAAGGATATAGAAATTAATGTCATGTGGCCC 4709
Qy      4666  AACAAATCTGAGGCTAATGAGCAGATTGCTTGTATCATCTGGGGGAGATGTCCAGC 4725
Db      4710  GTTGACAGGAGGCCAATGACCAAGTATGATGTATATCTTGTGAAGCATGAGCAGC 4769
Qy      4726  ATCCGCTCAATGCGCCAGTAGAGAGTCAAGAGCGTGTCTCACTCACACTTCCA 4785
Db      4770  ATTAGTCAAAATGCGCCGTGGAAGATCGGAAGGCTCCACACCACTAGACGCTGCT 4829
Qy      4786  TGCCTGTATTAATAGCTATAGCGGTGAGCGGCTATACAGTTGCGCTTGCAGAA 4845
Db      4830  TCTTGTGATCTCATGTATATATCTTCAGAAAGATACCAACCTTAAAGCTTCAGTCCA 4889
Qy      4846  GAACAGTTCGCGGATATGCTCATCTCTGTTGCGGAAGTACAGGATCACAGGCGTGAG 4905
Db      4890  GAACAAATTAATCTGTGTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4949
Qy      4906  AAGCTACAGTGCAGCAACCAAGCTCTGTTTCAGCGCTGTACCAACCGCTGTACACCCC 4965
Db      4950  AAGATTCAGTGTCTCCACCTATATCTGTTTCTACCTAAGTGTGCTGATCATTCACA 5009
Qy      4966  AGGAAGTAC 4974
Db      5010  CGGAAGTAC 5018

RESULT 13
VEUS5360      11420 bp      RNA      linear      VRL 12-JUL-1999
LOCUS      Venezuela equine encephalitis virus strain SH3, complete genome.
DEFINITION      U55360
ACCESSION      U55360
VERSION      U55360.2
KEYWORDS      GI:5442468
SOURCE      Venezuela equine encephalitis virus
ORGANISM      Venezuela equine encephalitis virus
Virus; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; VEEV complex.
REFERENCE      1 (bases 8366 to 10042)
AUTHORS      Weaver, S.C., Salas, R., Rico-Hesse, R., Ludwig, G.V., Oberste, M.S.,
Boshell, J., and Tesh, R.B.
Re-emergence of epidemic Venezuelan equine encephalomyelitis in
South America. VEE Study Group
Lancet 348 (9025), 436-440 (1996)
JOURNAL      Lancet
MEDLINE      96332169
PUBMED      8709783
TITLE      2 (bases 1 to 11420)
AUTHORS      Wang, B., Barrera, R., Boshell, J., Ferro, C., Freiler, J.E.,
Navarro, J.C., Salas, R., Vasquez, C., and Weaver, S.C.
Genetic and phenotypic changes accompanying the emergence of
epizootic subtype IC Venezuelan equine encephalitis viruses from an
epizootic subtype ID progenitor.
J. Virol. 73 (5), 4266-4271 (1999)
JOURNAL      J. Virol.
MEDLINE      99214366
PUBMED      10196323
REFERENCE      3 (bases 8366 to 10042)
AUTHORS      Weaver, S.C.

```

TITLE Direct Submission
JOURNAL Submitted (18-APR-1996) Pathology, Univ. TX Med. Branch, 301
REFERENCE University Ave., Galveston, TX 77555-0609, USA
AUTHORS 4 (bases 1 to 11420)
Wang, B. and Weaver, S.C.
JOURNAL Direct Submission
TITLE Submitted (12-JUL-1999) Pathology, Univ. TX Med. Branch, 301
University Ave., Galveston, TX 77555-0609, USA
REMARK Sequence update by submitter
COMMENT On Jul 12, 1999 this sequence version replaced gi:3849806.
FEATURES
source
1. 11420
/organism="Venezuelan equine encephalitis virus"
/mol_type="genomic RNA"
/strain="SH3"
/db_xref="taxon:11036"
45..7502
/note="contains nonstructural proteins 1, 2, 3 and 4;
in-frame termination codon is read through at low
frequency"
/codon_start=1
/transl_except=(pos:5658..5660,aa:OTHER)
/product="nonstructural polyprotein"
/protein_id="A043358.1"
/db_xref="GI:5442470"
/translation="MEKRVHVDIEDSPFLRALORSPOFEVEAKQVTDNDHNAARF
HLSKLIETEDPDTLIDIGAPARMYSKHKXHCIPMRCAEDPRLVYATKLRK
NCEITPEKEDKKKELAAVMSDPLDETETCKLDDSSREGQAVAYQVYVADPT
SLHQAOKGVRAVWIGDTTPPEMKLAGAYPSYSTMADEIVLIRANGLCSDDM
ELSRHGMILRKKYLRKSNVLFVSQSTIHEKDLIRSGFLRGKQNTCC
ETIYSCGVVYKRIASPGLYGKPSGVAATHRGSCFLCKVTDLNGEYSFVCTV
PATLCDQMTGLATDVADDAQKLVGNORIVNAGRTORNTMKYLLPVQAFA
RMAKEYEDODERPLGLRDLVMSGCMFRHKITSYRPTOTIKNSDHSF
VLRIGSTLEIGRTIRKLEBKRPSPILTADVOEAKCADEAKREAEAL
ALPLADVESEPLEADVDFMLQEGASVETPGLIKIVSYADEKIGSYAVSPDA
VLNEXKLSCHPLAEQVIVTHSGRKSAVEPHGVVVEGALIVQDFQALSEBA
TIVSNEREFVNRVLIHATGGLANTDEEYKIVKPSHEGEYVLDIRKQCVKESEV
TGSLGELVDPEPAVESLRTPEAPVQVPTIGYVAGSGSKGISATVTKDL
VLSAKKENCALIRIDVKMKGLDVNARTVSVILNGKHPETLYIDEAPFCHAGTIR
ALAIIRPKKAVLCGRKQCFPMNCKLKHFNHETCTQVPHKSISRCKTSYTSVS
TLFYDKMRITNPETIKIEDTGSTIKKODLLITCRGVNKOLOIDYKNEIMTA
ASQGLTRKGVAVYKVENPFLVAPTEHNVLLITREDRLVMTGLADPVIKTLTA
YPGNFIATIEEMQSAEIMHILIERPDTVFQNKANVCAKALVPLKTAGIDMTT
EOMNTVDFETDKASHAEIVNOLCVRFEGDLGFSFATVLSIRNNHMSPP
NMVGLKEVVRQLSRVRPLPRAVATGVMDNGALCNDPRLINLVARRLPHAVI
LHNHREHPOSDFSVSKLGRVTVVSEKLSVPKGTDMLSDQBEATPRALDGTIG
DVRYDVFIVNRTPYKHYHQCEDEHAKLSMLTKACLHNGGTCVSIQYVGR
ASESIIIGAIARQKFSRVCKRSLSETEVLFVPIGYDRKARTNPKLSTLTNYT
GSRHAEAGCAPSYHNVGDIATATEGVIINANSKGQGGGVCALYKKEPESDLP
IEVGKRLVKGAAKHIHAVGNPNKYSVEVDGKOLABAYSIKIVDNNYKVALP
LLSTGIPSGNKRDLTOSLHLLTALDITDVAVALYCRDKKEMTLKAEVAREVEEI
CISDSSVTEPDALVYVHPRKSLAGKRYSTSGKTPSYIEGRTFOAKDIAETNA
MIPVATANEQVCTYILGESMSTGRKCPVESEASTPSTPLCLCIHAMTEPVEQL
KASPEQITVCSFPLPKRYITGVQKIQCSPLIFSPKVPYIHPRYLVPTKVEEN
QSTEGTEPAIPVNDATRTMPEDIIIEEEDSISLSDPTQLOVEDIDIGPS
VSSSSIIIPASDFDVSLSITDLEIGTSYSEASAEIVTSFASMEFLAPAPAR
TVENPMPAPARTPTPLAPCRASSTISVTPGNVRIITREQLHLLTPRASRA
SKTSIVNPGVNRVITREPEARVAAQOQRFDAGATFSSDITQGLHQCSVQYTL
SEVUERTELEISTAPRDOKEKELRKQQLNTPANRSYQSRVENMALTARI
LOGGAYLKAEGVECYRTLAPVPLVSSVNRASSPVAVEAONAMLKEPFLAV
CIIEPDAYLMDVAGASCLDTPASPKASFPKSHSYPTIRSAVPSAIOQLN
VLAATKRNKNCVOTRELPUVLSAANVECKKACNNVEYETKEPDIRTEENIN
YITLKGPKAALFAKTHNLNMODIMPDRVMDLKRDPCTHPTKEERPKQVOI
AADPLATVYICGIRRELVRRLNVLIRIHTLTPMSADEPAITIAHFPQBDCTLED
IASFDKSEDDMALTLMLIEDLVADLFLIAAFGEISHLPTKXFRFAMK
SGMFLTFVNTVINIVIASVRLRRLTSPCAALIGDNIIVKGYSLADWACRTAT
NMEYKIIDAVVEGKAPYFCGGFICLDTVGTACRVADPLKTLFLKGLAADDD
RRRLHESSTRMNVGLPELCKAVESYETVGTSSIIYMATTLASSVSPSYLRGAP
ITLYG"
7538..11305
/note="contains capsid, E3, E2, 6K and E1 proteins"
/codon_start=1
/product="structural polyprotein"

/protein_id="AAC71999.2"
/db_xref="GI:5442469"
/translation="MFPPOPMYPMPNPPAPRRPMPPTDPLAMOVBLTSS
MANULTRQRORDAPREBPAPKPREBAPQOKRGQCKKKOKCKKAKKAKK
GNKKTKPKPKRQVWMLKLESDRTPFIMLEKINGYACVGGKLRPHAPVEKIND
VLAALKTKRASKYDLEADVDPQWNRADTFEYHEKINGYACVGGKLRPHAPVEKIND
GVAKGSDSGRIIDNOGRVVAIVLGVNESRSLSVNMNNGVYDELLAEAAKCPKRS
LVTMLCLAVTPPCAPICIDRKPELTAMSLVNMNNGVYDELLAEAAKCPKRS
STEBLEFYKLT/RPYNARCIRCAVGSCHSIIAEVSDHDDVGLQTSGLDSS
GNIKGRTPYDMHGTIEEPLHGVSLHSTPSCHLIDVGVFLARGCAGSITMEPK
NSTHSCSVYEYKPNVGRLEIITHPREHAEQCVYADDAORGAVEMHLPGSV
DLSVLSGSSVTVTPARTSALVECECGGTSETITNRKQSPQCKQCRVRLQ
NDKRVNSDCLKPRAAGATLKGKLVPELADGKTVLADEPMITTEGRFVSLHLK
NPYLTROLADPEPHYTHELISEPAVNFVTEKGEFWMGNHPRKFAQETPAGNP
HGLPHEVITHYHARPEWSTIGSICAIITVSAVASTLFEKRSVCLTPRLTPA
RMLCLAVICCARBARAETWESLDHMMNNOOMFOTLIPLAALIVTRILKVCV
VPEPLVAGAGAGATBHATMPSOACISINTIVNRAGVPLPISITPTKILPTYN
LEVYVCHYKTMGSPAIKCCGSOECTPTVPDQCFVGVYFPMGGAFCQDENT
QVKAAYMKSDDCLADIAEAYKATIASVQAFNITVGEHSIIWTVVNGEPTVNFANT
KLTAGPLSTAMPEDRKIVQYAGEIYVDEPEYAGQPGAFDIQSRTVSSDLVANT
NLVLOPKAGAIHVPTQAPSGPEOMKDKAPSLKFTAPRGCETITNPIAENCAVGS
ILPAFDLPDLFTRVSETPPLSAECTNHCYSSDQGGATVYKYSKSGKCAVHP
SGATLKEAVALTEQSGVTIHSTANIHFFRLQICTSVTCKGDHPREDHVTNP
OYHAQFTTAAVSTAMTWLTSLLGSSAVIIIGLVATTIVAMVTLNOKAN"

BASE COUNT 3220 a 2821 c 2895 g 2484 t
ORIGIN

Query Match 42.4%; Score 2119; DB 14; Length 11420;
Beet Local Similarity 64.5%; Pred. No. 0; Mismatches 1745; Indels 24; Gaps 3;
Matches 3220; Conservative 0;

10 ACTATGATCCATATATGAAGAATTCAGCTGACTTAAGTGTGACAGCCCTATGTC 69
30 AATTACTTACCCAAAATGGAGAAAGTTCAGCTTGACATCGAGAGAAGCAGCCATTCTC 89
70 AAGTGTACAGCGGACGTTTCCAAATTTGAGTTCAGAGCAAGCAGGTCACTGACAAT 129
90 AGAGCTTTGCAAGGAGACTTCCCGACATTTGAGGTAGAAACCAAGCAGGTCACTGATAAT 149
130 GACCATGCCAATGCCAGAGGCTTTGCCATGTGGCAACAAAGCTCATTTGAGAGGAGTGC 189
150 GACCATCTTAACGCAAGGCTTTGCCATGTGGCTTCAAAAGCTGATCGAAACGAGAGTGC 209
190 GACCGGACCAAGTATCTTGTGACATTTGAGAGTGGCGCCCTGACAGACATGACATTTCCAT 249
210 GACCATCCGACAGATCTTGTGACATTTGAGAGTGGCGCCCTGACAGATGATTTCTTAAG 269
250 CACCGTATCATTTGTATCTGCTCTATGATTAAGCGCTGAAGACCCGACAGACTACAAACGG 309
270 CATTAAGTATCATTTGATCTGCTCATGATGATGTCGGAAGACCCGACAGATTTGTACAG 329
310 TATTCGAAAGACCTTAAGAAAGT-----GACATTCGACAGAAAGACATAGCCTCTAG 363
330 TATTCGAACTAAGCTGAAGAAAGTAAATTCGAGAAATTAATCTGACAGAAATTCGAGAA 389
364 GCGGCAAGCTGCTGAAGTCAATGTCAACACAGACCCAGACAGACTCCATCTGTGTATG 423
390 ATGAAGAGCTTGGCGCGCTCATGAGGACCTTGACCTGGAATCTGAGATGATGTGCTC 449
424 CACACAGAGCCACGTGTAGGTACTTTGAGAGTGTACAGATTAACAAAGTGTATGCGCA 483
450 CACACAGAGCATATCTGCTGACAGAGGCAAGTCCGCTTTTACACAGATGTATACGCA 509
484 GTCCATGACAGCATCAATCTACACAGAGCGCTTAAAGAGTTAGACAATTTATCTGG 543
510 GTTGACGAGCGCAAGTCTCTTACCAAGCGCAGCAAGGAGAGTTAAGTGCCTACTGG 569
544 ATAGCTTTGACAGAGCCCTTTTATGTACAAAACATGCGAGGTTCTTACCTTACTTAC 603
570 ATAGGCTTTGACAGAGCCCTTTTATGTTAAGAACTTGGCTGAGATATCATCTATAC 629
604 AACACGAAGTGGCTGACAGAGAGTATGGAAGCAGTAACATGGCTTGGTGAATCA 663

Dh 630 TCTACCAATGGGCGGACGAAAACCGTGTAAACGGCTCGTAACATAGCCCTATGCAAGCTCC 689
Qy 664 GATCTTCAGSAGAGAGGCTTGAAAACTCTCAATCTCTTAGAGAAAGAGGCTCCAACT 723
Dh 690 GACGCTATGAGAGCGGTCACTGAGAGGATGTCCATCTTTAGSAAAGATTTGAAACCA 749
Qy 724 ACTAATAGATCAATTCGCGTTGGTTCAACATCTTACACAGAAAGATGATCACTGTTA 783
Dh 750 TCCAAATATGTCCTATCTCTGTGGCTGCACTATACACAGAAAGAGGAACTTACTG 809
Qy 784 CGTACTGCGCATCTTCCAAACGTTTCCCTTGAAAGAAAGTCTAATCTTACAGATAGA 843
Dh 810 AGGACCTGGCACCTCGCTGTATTTCACTTACGTGGCAACAAATTTACCATGTGCG 869
Qy 844 TGTGGGACCATGTGACGTGTGAGGGTACATCAATAAAGATAACATGACCCAGCA 903
Dh 870 TGTGAGACTATAGTTAGTTGCGAGGTAAGTGTAAAGAAATAGCTATACGTCCAGGC 929
Qy 904 CTATACGGTAAAGTTGAGAACTTGGCGTCCACAAATGCAATGCGAGGGTTCTTGAAGTTC 963
Dh 930 CTGTATGGGAAACCTTCAGGCTATGTGTCACGATGCAACGGGAGATCTTGTGCTGC 989
Qy 964 AAAGTCAAGATACGCTGCGCGGCGAGAGGGTTTCTTGTGTGTATGATATGCA 1023
Dh 990 AAAGTCAAGACACATTTGAACGGGAGAGGGTCTCTTTTCCGTGTGCACTATGTGCA 1049
Qy 1024 GCCACACTTGGCATCAGATGACAGGGATCTTGGCACTGAGGTAGTGGATGACGA 1083
Dh 1050 GCTACATTTGTGACCAAAATGCTGCAATCTGGCAACAGATGCTAGTGGAGCAACGGG 1109
Qy 1084 CAAAACTATTTGGTGGGCTCAACCAAGATTTGTCTCAATGTAGAGACGCAAAAT 1143
Dh 1110 CAAAACTGCTGGTGGGCTCAACAGGCAATGTCTCAACGCTGCAACCAAAAGAAAC 1169
Qy 1144 ACTAACAATATCAAACTATTTATTTACAGTGTGCGCCAGGCGTTTCCAGTGGCG 1203
Dh 1170 ACCAATACCATAAATATTTATTTTGGCCGTAGTGGCCAGGCAATTTGCTAGTGGCA 1229
Qy 1204 CGTGAACATCGTGCAGCTTGGACGACGAAAGAACTAGGGGTGGGGAGGCACTCTT 1263
Dh 1230 AAAGAAATTAAGAAAGATCAAGAAATGAGAGGCCATGAGACTACGAAATGACAGTTA 1289
Qy 1264 ACTATGGGCTGCTGCTGGGCTTTCAAGCCAGAAATTCATCTATCAAGAAAGCT 1323
Dh 1290 GTTATGGGGTGTGCTGGGCTTTTAGAAGGACAAAGATTAACATATTTATTAAGGCCA 1349
Qy 1324 GGTACGCAAACTTAAGAAAGTACCTGCGCTTTGACTCATTTGTGATTTCCAGCCTT 1383
Dh 1350 GATACCCAAACCATCATCAAGTGAACAGCGATTTCCACTCATTTGTGCTGCCAGATA 1409
Qy 1384 ACCAGCCACGGGCTGATATGCGGCTTCCGCCGTACAGCTGCTGTGAACCACT 1443
Dh 1410 GGCAGTAAACATTTGGAAGTCCGGCTGAGAACAAATTCAGAAATGCTAAGAAAGCGC 1469
Qy 1444 GTCAAAACCCGACCGCTATTTACATGCGCGATGTGAGACTCTCGTGGCTTACAGCA 1503
Dh 1470 AAGAGCCGTCACCTCTCATTTACTGCCAGACGATACAAAGAGCTAAGTGGCCAGCAT 1529
Qy 1504 GAACTGAAGAAAGTGGCTGACGCGGAAGATCAAGAAAGCCCTGCCACCTTGTCTCT 1563
Dh 1530 GAGGCTAAGGAGTCCGTGAAGCCGAGGAGTTCGCGCTGCACTCAACCTTTGGCAGCA 1589
Qy 1564 GAAATAGAAAAGAACCGTGAAGGACAGAAAGTAACTCATATGCAAGAGGCAAGAGCA 1623
Dh 1590 GATGTTAGAGACCCACTCTGGAAGCCGATGTGACTTATGCTAACAAAGGCTGGGCGC 1649
Qy 1624 GGTACGCTGAGACACACGAGAGACATCAAGGTGACAAAGTTACCCAGCAAGAG 1683
Dh 1650 GGCCTAGTGAAGACCTCTGTGGCTTGAATTAAGGTTTCAAGCTATGCGCGAGAGCAAG 1709
Qy 1684 ATTGGCTCTTACGCTATTTCTACCCAGGCGGATTTGAATGTAAGAAACTGGCGTGT 1743
Dh 1710 ATTGGCTCTTACGAGTGTCTTCTCACAGGCTGTACTCAAGAGTGAAGAACTATCTTGC 1769

Qy 1744 ATCCACCATTTGGCGGAACAAGTACTGGTAAATGACTCAAAAGTATGGGACGAGATAC 1803
Dh 1770 ATTCACTCTCTCGTGAACAAGTCAATAGTATTAACACTCTGCGCGCAAAAGGCGTTAT 1829
Qy 1804 AAAGTGAAGCATACCAAGTAAAGTCAATTTGACCAAGAGGACGGCGCTCTGTCA 1863
Dh 1830 GCCGTGAACCTTACATGAAAAAGTATGGTCCAGAGGGACATGCAATACCTCCGACG 1889
Qy 1864 GACTTCAGGCAATTGAGGAGAGCGCTACGATCCGTTTCAACGAGAGGAGTTCTTAAAC 1923
Dh 1890 GACTTCAAGCTCTGAGTGAAGGTGCACATTTGTATCAAGAACGAGAGTCTGTAAAC 1949
Qy 1924 AGATCTGACACACATGCAATCAACGAGAGGCGCTTAAACATGACGAAAGTACTAT 1983
Dh 1950 AGGTATTTGACCAATTTATGTCACACATGAGAGGCGCTGAACACAGATGAATATTTAC 2009
Qy 1984 AAGACTGTAAAGACTCAGAGACACAGCTCAGAAATCGTCTTCGATTTGACGACGAAAG 2043
Dh 2010 AAAATTTGCAAGCCAGGAGACACAGCGGCAATCTGTATGATGACATGACGAAACAG 2069
Qy 2044 TGTGTTAAGCAGAAAGACGAGTCCCTTGTGCTTAAACCGGTATCTGTAGATCCACCA 2103
Dh 2070 TGGCTCAAGAAAGAACTAGTCACTGGGCTAGGGCTTACAGGCGAGCTGTGATCTTCC 2129
Qy 2104 TTTCAAGATTTGGCTACGAGAGTCTCAAGACAGAGCAGACGCTCACAAAGTTCCA 2163
Dh 2130 TTCAATGATTTGCTTACGAGAGCTTGAAACACAGCGGCGCTCTTACCAATGACCA 2189
Qy 2164 ACCATCGAGTCTATGAGTGTGCAAGTTCAGGTAAATCTGAAATCATCAAAAGCCTGTG 2223
Dh 2190 ACCATAGGGGTGTATGGGTGCGCGGTTCAAGCAAGTCTGGATATTTAAAGGCCAAGT 2249
Qy 2224 ACTAAGAAAGATCTGTGTGTAGTGCAGAAAGGAAAACTGGCAGAAATCATCGGAT 2283
Dh 2250 ACCAAAAAGATCTGTGTGTAGTGCAGCAAGAAAGAACTGGCAGAAATATAAGGAGAC 2309
Qy 2284 GTAAGAGATGAGACGATGGAATGTCTGTAGAGCTGTGATGATTCAGTGTCTTAAT 2343
Dh 2310 GTCAAGAAATGAAGAGGCTGAGCTCAATGCGAATCTGTGATCTCATGCTCTTGAAT 2369
Qy 2344 GGGGTTAAGCAACCCGTTTAACTCTGTATCAATGTAAGGCAATTTGCTGCCATCAGGG 2403
Dh 2370 GGATGCAAAACCCGTGAGAGCCCTGTATTTGACGAAGCTTTGCTTGTATGACAGGC 2429
Qy 2404 AGCTGTGCGACATGATTTGCAATGCTCAACCTTAAAGAAAGTGTATTTGTGGGGACCA 2463
Dh 2430 ACTCTCAGAGGCTCATAGCATATTAAGACCTTAAAGGACGATGCTCTGGCGGATCCA 2489
Qy 2464 AAACAAATGCGGCTTTTAAATGATGATGAGCTGAAAGTAAATTTAAACATGACATGAC 2523
Dh 2490 AAACAGTCCGGCTTTTCAATATGATGTCCTGAAAGTGAATTTTAAACACAGATTTGC 2549
Qy 2524 ACTGAAGTATCAATTAAGAGATCTCTGAGAGTGCACACAGACTGTAAACGCCATCTGC 2583
Dh 2550 AAGCAGGCTTTCACAAAGAGATCTCTGTGTGTCATTAATCTGTGATCTTCCGCTGCT 2609
Qy 2584 TCCAGGCTCTTACAGACAGAGTGAAGACGTTAACCATGTGCTGATTAATATC 2643
Dh 2610 TCAACCTGTTTTCAGCAAAAGATGAGACGAGAACCCGAAAGAGACTTAAGATTAG 2669
Qy 2644 ATGATATACAGAGGACACAAAGCCGACAAAGATGATGATTTCAATCTGTTTTCAG 2703
Dh 2670 ATGACACTACTGGAGTAAACCAAGAGAGAGTCAATCTCATCTTGTTCAG 2729
Qy 2704 GGAATGGTGAACAGCTACAGATTGACTCAAAATATCGAAATCATGATCGCGCTCA 2763
Dh 2730 GGGTGGTAAAGAGGTTGCAATATGATTAACAAAGCAACGAAATATACGCGAGCTGCC 2789
Qy 2764 TCGAAGGACTTACGCGGAAAGCGTTTATCTGTCAAGTCAAAAGTCAAGAGATCCA 2823
Dh 2790 TCTCAAGGCTGACCCGTTAAAGGCGTGTATGCTGCTGTCAAGAGTGAAGAAATCTCC 2849

QY 2824 CTCTACTGCGAGACTTCTGAGCAGCTGMAAGCTGTACTTACTTACGACGACAGAAAAAGCATT 2883
DB 2850 CTGTACCGCAACCCACTCTGAGAACATGTGAACTCTACTGACCCGTTACGAGAGACCGTATC 2909
QY 2884 GTCTGGAAGACGCTAGCTGTGTATCCCTGATTAAGACACTTACAGCTTAATATCCGGG 2943
DB 2910 GTGTGAAAAACACTAGCCGCGGATCCATGTGATTAACAACTGACGGCCAAATATCTGGG 2969
QY 2944 GATTTCACGGCTTATTTGAGACACTGAGAGCGGCAACGACGCACTTATATGGACCGCTT 3003
DB 2970 AATTTCACCTGCAAGATGAGAAATGCGAACAGACATGAGGCACTATATAGGCACTC 3029
QY 3004 CTGTATAGCCGCGAGACAGCTGATGTGTTCCAGATTAAGTGAAGCTCTGCGGCGAG 3063
DB 3030 TTGGAAGACCCGACCTTACCGAGCTATTTCCAGATTAAGGCGAACGTGTGTGGGCCAAG 3089
QY 3064 GCTTTAGACCCAGTCTTGCCCAAGGCCAACATTTGCTGACAGACAGCACTGGGAGAGC 3123
DB 3090 GCTTTGGTCCGCTGACTGAAACCTGACAGGCACTAGACCACTGAACCAATGGAACT 3149
QY 3124 TTGCAACCATTTCAAGCATGAGACGAGCTCACTCTGAAAATGGCACTGAACTTTCTTTC 3183
DB 3150 GTGATTAATTTGCAAGCGAACAGCTCACTCAGCAGATGATGTAAGAACCAACTATGTC 3209
QY 3184 ACCAGTTCTTTGAGATGACCTGGAAGTGGTATTTTCCGCTCTACCGTGCACCTT 3243
DB 3210 GTTAGTCTTTTGGACTGACCTGGACTCCGGTCTATTTTCTGACCACTGTTCCGTTA 3269
QY 3244 ACTTACAGGATTCAGCACTGGGATTACTGCGCAGGAGAAACATGTATGGCTTAAATGA 3303
DB 3270 TCCATTAGGAATATCACTGGGATTAATCCCGTCCCTTAATGATGATGCGGCTGAATPAA 3329
QY 3304 GAGGTACGAAGAGTGTGTCAGCGGCAATCCGTCATCAAGAACGCTTGAACAGGCG 3353
DB 3330 GAATGTGCTCGTCACTCTCCGCAAGTACCCAGCACTGCTCGAGAGTGCCTATCTGGA 3389
QY 3364 AGGTAAGCTGATATTAAGGAATATATCAATCAAGACTACTCTCAACCAATTAATGTGTT 3423
DB 3390 AGAATCTATGACATGAACACCGGCGCACTGCAATTAATGATCCGCGATTAATCTAGTA 3449
QY 3424 CCATTAATGCGCGGTGCCCCACTCGTTGATCGTTGACCAAGAGCAGGAGTAACT 3483
DB 3450 CCAATGAAACAGAAAGCTGCTCATGCTTTAGCTCCACCATTAATGAACCCACAGAGT 3509
QY 3484 GATACACGCGATTCCTATCTAAGATGAAGGCAAACTGTGTGTGATGAGGAGATCCT 3543
DB 3510 GACTTTTCTTCAATTCGTAGCAAACTGAAGGCGAAGCTGTGTGTGTGCGGAGAGG 3569
QY 3544 ATCAGCAATTCAGGGAAGAAAGTAGAGTCCATGGGTCCAATGGCCATTAATACCATCAG 3603
DB 3570 TTGTCCCTCCAGGCAAACTGTGTGATGCTGTGTGACAGCCAGCTTGAGGCTTACTTGA 3629
QY 3604 TGTGATCTCATTTTGGGAATACCTAGCCATGTGCGTAATATGAATTAATCTTTGTCAAT 3663
DB 3630 GCTGCGCTGATTAAGTATCCAGGTGACGTGCCAAATATGACATTTGATTAATTAAT 3689
QY 3664 GTTAGAACCCCGTACAGAAACCATCACTACCAAGTGGAGATACAGCTATCCACAC 3723
DB 3690 GTGAGGACTCATTAATTAATCAATCATTAATGAGAGTGAAGACCAAGCCATCAAGTTT 3749
QY 3724 AGCATGCTAAGCTTAAGGCTGTCCACCACTGAACATGCGGAAACATGTGTGGCTATA 3783
DB 3750 AGCATGTGACCAAGAAAGCTGTGTGTGATTTGAATCCGCGGAACTTGCTGACGATA 3809
QY 3784 GGGTATGGGCTTGTGTATCGGCAACCGAGAAATATCATCTGCGGTGACAGCTCATTT 3843
DB 3810 GGTATGTTATCGCTGACAGGGCCAGGAAAGCAATTAATGTGTGATATAGCCGCAAGTTC 3869
QY 3844 AGGTTTACCCGTGTCTGTGACCTTAAGAAACCTGCCAAATATCTGAGGTCTCTTGTG 3903
DB 3870 AAGTTCCTCCGCGGTATGCAAAACGAAATCTCACTTGAAGGAGAGAGTCTTGTGTGTA 3929
QY 3904 TTCTTCGGAAGAGCAACGCGCAACCAACATGACCGAGACAGACTCGGTGTGTGCTT 3963

DB 3930 TTGATTTGGTACGATCCGAAAGCCCGTACGCAAACTCTTAACAAGCTTTCATCTACCTTG 3989
QY 3964 GACAACTCTATCAAGGCTCAACAGGTACGAGCAGGAGAGCTCCAGGCTACAGAGTG 4023
DB 3990 ACCAACTCTATACAGGTTCCAGCTTCCAGAAAGCCGAGATGCCACCTCATATCATGTG 4049
QY 4024 ATCAGAGTGAATTAACAGAGGCTGACCAAGCTATGCTTAATGCTGTAATAGCAAA 4083
DB 4050 GTGCGAGGAGCAATTTGCAAGGCTCACGAGAGTGTATCAATATGCTGTAACAGCAAA 4109
QY 4084 GGTCAACAGGTTCCGAGTGTGGGTGACCTGTACCGAAATATGCGCGCTCTTTGAT 4143
DB 4110 GACAAACCTGCGGAGGGGTGTGGAGGCTGTATTAAGAAATTTCCGGAAGAGCTTGAT 4169
QY 4144 AGACAGCAATAGCTGTGCGGAGCGGCTGACCTGTGAAGACCAAC---GCTCATATA 4200
DB 4170 TTACAGCCGATCGAAGTGGGAAAAAGCCGACTGTCTAAAGTGTGAGCTTAAGCATATCAT 4229
QY 4201 CATGCTGTAGGACCAATTTTCTTAAGATGCGGAAACCGGAGGCGACTTAAAGCTGCA 4260
DB 4230 CATGCGTAGGCGCCCACTTCAACAAAGTTTCGAGGTGAGGTGCAAAACAGTTGGA 4289
QY 4261 GCTGCTTACATGAGCATAGGCTCATGTCACAGCTGAGCGGATTAACAAAATATCAATA 4320
DB 4290 GAGGCTTATAGTCCATCGTAAATTTGTCAACGATTAACAATTAAGTCAAGTACGATT 4349
QY 4321 CCGCTACTGTCAACCGGCACTATTTCTGTGTGCAAAAGATGAGTATGCAATCTTGCAT 4380
DB 4350 CCACGTGTTGTCACCGGCACTTTTTCGGGAAACAAATATGACTTAACCAATCAATTAAC 4409
QY 4381 CACCTGTTCACTGCTTTCGACACTAGCGGATGCGGATGACCAATATTTCTTGTGATPAA 4440
DB 4410 CATTTGCTGACAGCTTTAGACCACTGATGCAATGTATACCATATCTGACAGGACAG 4469
QY 4441 CAATGGAGACACAGATATGAGGCCATTACCGCAAGAAAGCGTGAATTTCTGAT 4500
DB 4470 AATGTGAATATGACTCTCAAGAAAGCAGTGGTATGAGAGAAAGCAAGTGAAGATATGT 4529
QY 4501 GATGACAAAGC-----AGTAGAATTTGACTTTGTGTGACAGGTCACCCCAAC 4545
DB 4530 ATATCCGACACTTTGCGTGAAGAACCGGATGCAAGCTGTGAGAGTACATCCGAA 4589
QY 4546 AGCTCTTTGGCAGCAGACAGGTTACTCCGTCAATGAGGCAAGTTGATTAATCACTG 4605
DB 4590 AGTTCTTGGCTGGAAGAAAGGCTTACAGACAAAGCATGCAAGACTTCTCATATTTG 4649
QY 4606 GAAGGTACACGATTCATCAGACCGCCAGAGACATTTGCCAAATCCATGCAATGTGCC 4665
DB 4650 GAAGGACCAAAATTTTATGAGGGCGGCAAGGATATAGCAAAATTAATGCAATGTGCC 4709
QY 4666 AATCAATCTGAGGCTAATGACAGATTTGCTTTGTAATCTGTGGGGAGATATGTCCAGC 4725
DB 4710 GTTGCAACGAGGCGCAATGAGCAAGTATGATGATATCTTTGTGTAACATGAGAGC 4769
QY 4726 ATCCGCTCAAAATGCCAGTATGAGAGTCAAGAGCGTCTGCTCACTCACTCACTTCA 4785
DB 4770 ATTAGTCAAAATGCCCCGTGAAAGATCGAAGCTCTGACACCACTTACAGCGCTCCT 4829
QY 4786 TGCCTGTATTAATACGCTATGACGCTGAGCGGCTATACAGTTGCGCTTGCAGAA 4845
DB 4830 TGCTGTGATTCATGCTATGATGCTCAGAAAGAGTACAAAGCCTCAAAAGCCTCAGTCA 4889
QY 4846 GAACAGTTCGCGGTATGCTATCTTCTGTGCGAAAGTACAGGATCAACAGCGGTGACAG 4905
DB 4890 GAACAAATTAATCTGTGCTCATCTTTCATTTGCGCAAGTATGAAATTAATCTGTGTGACAG 4949
QY 4906 AAGCTACAGTGAAGCAACAGTCTGTTTCAAGCGTCTGATCAACAGCGGTGATCAACCC 4965
DB 4950 AAGATCCAGTGTCCAGGCTATATGTTCTCACTTAAGGTGCTGCTGATCATTTCA 5009
QY 4966 AGGAAGTAC 4974

DB	5010	CGAAGTAC	5018
RESULT 14	AF375051	11423 bp	RNA
LOCUS	AF375051		linear
DEFINITION	Venezuelan equine encephalitis virus nonstructural polyprotein and structural polyprotein genes, complete cds.		
ACCESSION	AF375051		
VERSION	AF375051.1		
KEYWORDS			
SOURCE			
ORGANISM	Venezuelan equine encephalitis virus		
REFERENCE	Venezuelan equine encephalitis virus		
AUTHORS	Alphavirus; VEEV complex.		
TITLE	1 (bases 1 to 11423)		
JOURNAL	Brault, A.C., Powers, A.M., Medina, G., Wang, E., Kang, W., Salas, R.A.,		
MEDLINE	De Siger, J., and Weaver, S.C.		
PUBMED	Potential sources of the 1995 Venezuelan equine encephalitis		
REFERENCE	subtype 1C epidemic		
AUTHORS	J. Virol. 75 (13), 5823-5832 (2001)		
TITLE	2 (bases 1 to 11423)		
JOURNAL	Weaver, S.C.		
MEDLINE	Direct Submission		
PUBMED	Submitted (30-APR-2001) Pathology, Univ. Texas Medical Branch, 301		
REFERENCE	University Blvd., Galveston, TX 77555-0609, USA		
AUTHORS	Location/Qualifiers		
TITLE	1. 11423		
JOURNAL	/organism="Venezuelan equine encephalitis virus"		
MEDLINE	/mol_type="genomic RNA"		
PUBMED	/strain="P676"		
REFERENCE	/db_xref="taxon:11036"		
AUTHORS	/note="isolated from mouse brain antigen"		
TITLE	25. 7506		
JOURNAL	/note="nsP1-4"		
MEDLINE	/cdate="1998-01-01"		
PUBMED	/product="nonstructural polyprotein"		
REFERENCE	/protein_id="AAK66989.1"		
AUTHORS	/db_xref="GI:14549693"		
TITLE	/translation="MEKRVHVDIEDSPFLRALQNSFPQFEVAKQTDNDHANARAFS		
JOURNAL	HLASLTLEVDPSDTLLDIGSAPRRMYSGKHCICIPMCACEDPDLKYATKLK		
MEDLINE	NCKETDLEKMKKEALVMSDPLELETMCILHDBSCREGVAVQDVAVDGP		
PUBMED	SLYHQAQGVRAVWIGEDTTPFMKTLNAGVAPSTMAWETVLTANIGLCSGDV		
REFERENCE	ERSRSMILIRKKYLPKSNVYLSVSGITIEHKRDLASMLPSYFHLRGKONTCCR		
AUTHORS	ETIYSCDSYVVRILNISPGYIGKPSGYAATMRBSFLCCRTDTLNGRVSPFCTY		
TITLE	PATLDQNTGLIADVSADDDKGLVGLNRI VNGKRTQNTNKNLPLVVAQAE		
JOURNAL	RMWKEKQDQDERPLRLDRQLVNGCCWAFRRHKTITIKRPDTOTIIKNSDFHS		
MEDLINE	VLPRGNSLTLEIGLTRIRKMLEEKESPLITAEIDIOACDAEAEVLEELR		
PUBMED	ALPLPAADFEETLEADVDMLOEAGASVEPRGLIKVTSAGEDKIGSAVLSQ		
REFERENCE	VKSKSLCSCHPLAEQVIVTHSGRKGAVAEPRKGVVPEGHAIPQDQVALSES		
AUTHORS	TITVNERFVNRYLHLIATHGALNTDSEYKYTVPSHDEYLYIDIRKQCYKEL		
TITLE	TGLGTGELVDPPEHFAVESLRTRPAAPVQPTTIGVGVSGSGSIIKSAVTKDI		
JOURNAL	VSVAKGECACIEIRDVKKMKGLDVNARIVDSVNLGCHPVEYTIIDAFACHGT		
MEDLINE	ALIIIRPKKATLCPDPRQCGFNNMCKLHFNHIEICQVHKISIRCTGSVTSVS		
PUBMED	TLFDRKMRITNPKETKLEIDTGTGTRKODLLITCRGWMKQLOIDRYKSNMTA		
REFERENCE	ASOGITRGVAVRVKYNENPLVPTSHVNAVLTREDRILVMTLTADDPKILTL		
AUTHORS	YQNTATITIEKQAEHDAIMRHILIERPPTVFOKNAVCACAKLVPLTKTAGIDMT		
TITLE	EOMNTVATFETDKASAEIIVNLQLCVRFGDLDSGLSAPLVPLISINNMMSPS		
JOURNAL	MYMGANKENVQLSRRYPOLPRAVATGVVMNMTGTLNAYPRINLVNRRILPAL		
MEDLINE	LHNEHPOSDESSFVSKLGRVTVLVGKLSVGMKQVMLSDQPAETRALDGI		
PUBMED	DVPKDIDIVINVRTPYKHHYQOCEDHAIKLSMLTKKACHLNPQGTGVSIGYAD		
REFERENCE	ASESTIGAIARQKRSRYCKKSSHEEVEVFLVETGCRKARTNPVLTSTINIV		
AUTHORS	GSRLHACGASRYVRDITATITGCVINANKSGQGGGVCALYKFKPESRDLQ		
TITLE	LEVGARLVKQVTKIITHAVGENFKVSDEGDKLQVAYESIAITVNDNNYSVA		
JOURNAL	LSTIIFSGNKDRLTQSNHLITADTADVADVACRDKMERTLKEAVARAEVEE		
MEDLINE	CISDSSVTEEDPAELVRHPRKSLAGKVSISDCKTSYLETGCFHAAOIDE		
PUBMED	MPVATENBOVCWYIIGESMSSISKCPVSESSSTPSTPLCICHAMPBRO		
REFERENCE	KASREOITVCSPLPKYRITLQKOCOSPIILSPVPVPIHRYKVIPTPES		
AUTHORS	PSRHQNSCTECPQPLVNVADRTTMEPEIITIEEESISILSGPFIQVQV		
TITLE	ADHSPSVSSSSWSIPIHASPDDVDSILTLDGASVTSVAENLNSYFARSEN		

Query Match	Best Local Similarity	64.5%: Pred. No. 0:	Matches 3223: Conservative	0: Mismatches 1750:	Indels 24:	Gaps 3
QY	2	CCCTACAAACTTAATGCATGCATTAATGGAAGAAATTCACGTTGACTTAGATGCTGACAGCC	61			
DB	2	CCCAAAACCAATTACTCTACCCCAAAATGGAAGAAATTCAGTTGACATCGAGGAAGACAGCC	61			
QY	62	CGTATGCAACTCGTTTACAGCGGACGTTTCCAACTTTGGAATGGAAGCAAGCAGCGTCA	121			
DB	62	CATTCTCTAGAGCTTTTCAACGAGGCTTCCGCGAGTTTGAGTGAAGACCAAGCAGGTCA	121			
QY	122	CTGCAATATGACCATGACCAATTCGACAGCGCTTTTGGCATGTGGCAACAAGCTCATTTGAGA	181			
DB	122	CTGATATATGACCAATCTATATCCGAGAGCGTTTGGCATCTGGCTTCAAAATGATATCGAAA	181			
QY	182	GCGAAGTCGACCGGGAACCAAGTTATCTTTGACATTGGAAGTGCAGCCGCTGACACATGCAC	241			
DB	182	CGGAGGTGGAACCATTCGACACAGATCCTTGACATTGGAAGTGGCGCCCGCAGAGATGT	241			
QY	242	ATTCCATATCACCGGCTATCTGTATCTTGCCTTATGATTAAGCGCTGGAAGACCCCGACACAC	301			
DB	242	ATTCTTAAGCATATAGATCATTTGATCTGTCTCGATGAGATATGTCGGMAATCCGACAGAT	301			
QY	302	TACAACGGTATGACGAAGAATTAAAGAAAAT-----GACATTACGACACAAGACATAG	355			
DB	302	TGTACAAATATGACACTTAAGCTGAAGAAAATTGCGAAGAAATTAATCTGACAGAAATTGG	361			
QY	356	CCTCTAAGGCGGACAGACTGCTGAGATGATGTACACACGACGACAGACCTCATCTC	415			

Db 362 ACAAGAAATGAAAGAGCTGCGCCCGTCATGAGCGAACCGCAGCTGGAATGAGACTA 421
Qy 416 TGTGTATGACACAGAGCGCCAGTGTAGATCTTTGGAGTGTAGCATATCAAGATG 475
Db 422 TGTGCTCCAGCAGATGATCATGTGCTGTGAGGGGCAAGTGCCTGTTTACAGAGATG 481
Qy 476 TGTACGAGTCCATGCAACCGCATCATCTACCAAGGGGCTTAAAGGAGTATGAGCAA 535
Db 482 TATAGCGAGTGTAGCAGACCGCAGAGTCTTATCACCAAGCCAAAGGAGTATGAGTCG 541
Qy 536 TTTATGAGTAGGCTTTGACACGACCCCTTTATGTACAAATAAGATGGCAGTTCTTACC 595
Db 542 CCTACTGGATAGGCTTTGACACACCCCTTTATGTATTAAGAATTGGCTGAGCATATC 601
Qy 596 CTACTTACACAGCACTGGGCTGACGAGAGTATTTGAAAGCAGTAACTTTGGCTCG 655
Db 602 CATCATCTCTACCAACTGGGCGCAGCAACCGTGTAAACGGCTGTAACTAAGGCTTAT 661
Qy 656 GTACTCAGATCTTCAAGSAGCAGGGCTTGGAAAACCTCAATCCTTATGAAAGAGAGGC 715
Db 662 GCAGCTCCGAGCTATGAGCGGTCACTAGAGGAGTCCATTTCTTAAAGAAATGATT 721
Qy 716 TCCAACTACTAATTAAGATCATATTTCTGCTGTGTTCAACATCTACACAGAAATAGAT 775
Db 722 TGAACCATCCAAATATGTCTTATCTGTGTGGCTGACATCTACACAGAAAGAGG 781
Qy 776 CACTGTATAGTACTGGCATTTTCCAAACGTTTCCACTTGAAGAAAGTCTTAACTTCA 835
Db 782 ACTTACTGAGAGGTGGCACCCTGCTGTATTTTCACTTACGTGGCAAGCAAAATTTACA 841
Qy 836 CAGGTATATGTGGAGCCATTGTCAGCTGTGAAGGTTAGTCATCAAAAAGATACGATCA 895
Db 842 CATGTGCTGTGAGCTATATGATGTGCAAGGTTAGTGTATTAAGAAATACCTATCA 901
Qy 896 GCCCAGCTATTAAGGTTAAAGTTGAGAACTTGGCGTCCACATGCATGCGAGGGTTTCT 955
Db 902 GTCCAGGCTGTATGGGAAGCTTCAAGCTATGTGCTAGCATGCAACGCAAGGATTTCT 961
Qy 956 TGAAGTGAAGTACAGATAGCTGCGCGCGCAGAGGGTTTCTTTTCTGTGTAGCT 1015
Db 962 TGTCTGCAAGGTACAGACATTTGAACGGGAGAGGGTCTCTTTTCCCTGTGCACT 1021
Qy 1016 ATGTACCAAGCCACCTTGGCATGATGAGGAGTTCTGGCAACTGACGTTATGTGG 1075
Db 1022 ATGTGCAAGCTTACTTGTGTACCAATGACTTGGCATCTGGCAACAGATGTCAATGGG 1081
Qy 1076 ATGACGCAAAAATCTATTTGGTGGCTCAACAAAGGATTTGCTCATATGTAGAGC 1135
Db 1082 ACGACGGGCAAAAATCTGTGGTGGCTCAACAGCGCATGTGTCAACGTTGCAACCC 1141
Qy 1136 AAAAATACTTACACAAATGCAAACTATCTATTAACAGTGTGCCCAAGGCTTTTCCA 1195
Db 1142 AAAAATACTTACACAAATGCAAAATTTCTTTTGTGCTGCTGCTGCTGCTGCTGCTA 1201
Qy 1196 GGTGGGGCGTGAACATGTGCCGCTTGAAGCAAGAAAGAACTAGGGGTGGGGAGC 1255
Db 1202 GGTGGGCAAGGAATATAGAGATCAAGAAATGAGAGGCCATAGGACTACAGATTA 1261
Qy 1256 GCACCTTACTTACTAGGCTGTGCTGGGCTTTCAAGACCCAGAAAATCATCTCACTTACA 1315
Db 1262 GACAGTTAGTATGAGGGGTGTGCTGGGCTTTTGAAGGACACAAATATCTATTTATA 1321
Qy 1316 AGAAGCTGTGTACGCAAACTTAAAGAAAGTACTGCTGCTTGTGATCTATTTGTGATTC 1375
Db 1322 AGCGCCCGATATCCAAACCATCTCAAGATGAACAGCGATTTTCACTATTTGCTGTC 1381
Qy 1376 CACGCTTACAGCAGCGGCTCATATGGGCTTCCGCGGTAGGCTCAAGCTGCTGCTTG 1435
Db 1382 CGAGATGTGAGATTAACACATGAGATCGGGCTGAGAAAGAAATGAGCTAG 1441
Qy 1436 AACCAATGTCAAAACCCGACCGGCTATTTACAAATGGCGATGTGAGAGATCTGCGTGGCT 1495

Db 1442 AAGAGCAAGAGACCGCTCACCTCTCATTTACTGCCGAGCATATCAAGAGCTTAAGTCG 1501
Qy 1496 TACAGCAAGAAAGTGAAGATGTGCTGCAGCGGAAGAGATCAGAAAGCCCTGCAACCT 1555
Db 1502 CAGCGATGAGGCTTAAGAAATGCTGTAAAGCCAGAGAGCTGCGCTGCATACCACTT 1561
Qy 1556 TGTCTCCCTGAATATAGAAAAGAGACCGTGAAGGACAGATAGACCTTATATGCAAGAG 1615
Db 1562 TGGCAGCTATTTTGAAGAGACCCACTGTGAACCCATGTGTGACTTGTATCAAGAGG 1621
Qy 1616 CAGAGCAGGTAGCGTGTAGAGACACACAGAGACATCAGGGTGAACAATTTACCAGGCG 1675
Db 1622 CTGGGCGCGGCTCAGTGTAGACACTCTGTGGCTTTGAATAAAGTTTACCACTATGCCGCG 1681
Qy 1676 AAGAAAGATTTGGTCTTATGCTATCTTACCCCAAGGGGATTTGAATAGTGAAGAAC 1735
Db 1682 AGACAAAGATGTGCTTTAGCAGGTGTTTCTCCACAGGCTGTATCTCAAGATGTAGAAC 1741
Qy 1736 TGGGCTGTATCCACCCATTTGGCGGAACAAGTACTGTATATGACTCACAAAGGTAGGGCAG 1795
Db 1742 TATTTGCAATTCACCTCTGCTGTAACAAATCATATGTATACACACTTGGCGGAAG 1801
Qy 1796 GAGATATCAAAAGTGCAGCCATACCAAGGTATAGTCAATTTGAACAAAGGACGGCTCC 1855
Db 1802 GCGGTATGCGGTGAACCTTACCATGGAAGATGTGTGCCAGAGAGGACATGCAATAC 1861
Qy 1856 CTGTTCAGACTTCCAGGCAATGATGAGAGCCCTACAGTCTTTTCAAGAGAGGAGT 1915
Db 1862 CCGTCCAGACTTTCAAGCTCTGATGAAGTGTCCACATCTGTATCAAGAACGAAAGT 1921
Qy 1916 TCGTAAACAGATACCTGACCAATCCATCAACGAGAGGCGTAAACATGTACAGAG 1975
Db 1922 TCGTAAACAGATACCTGACCAATTTGCTACACATGTAGAGAGCGCTGAACAGATGAG 1981
Qy 1976 AGTATTAAGACTTAAAGACTTACAGACACAGCTCAAGATACGTCTTGAATTTAGC 2035
Db 1982 AATATTAACAACTGTCAAGCCACGACGACGCGCAATACCTGTATGACATTCACA 2041
Qy 2036 CAGAAAGTGTATTAAGCAGAAAGCAGAGTCCCTTGTGCTTAAACCGGTATCTGTAG 2095
Db 2042 GAAACAAATGCTCAAAAGAAATTAATCACTGGGCTTGAAGGCTTACAGGAGTGTGG 2101
Qy 2096 ATCCACATTTCAAGATTTGCGTACGAGAGTCTCAAGACGACGACGACCTTACA 2155
Db 2102 ATCTTCCCTTCCATGATTTTGCTTACGAGAGTGTGAACAAGCTCGGCGCTCTTAA 2161
Qy 2156 AAGTCCCAACCATGAGCTATGAGTGTCCAGGTTCAAGTAACTTGAATCTCAAAA 2215
Db 2162 AAGTACCAACCATGAGGCTGTATGTGCGGGGTGAGGCAAGTCTGGCATCTTAAAA 2221
Qy 2216 GCGGTGTACTTAAAGAAATCTGTTGTGAGTGTGAAGAAAGAACTGGGCAAGAAATCA 2275
Db 2222 GCGGAGTCCAAAAGAAATGTGTGTGAGCGCAAGAAAGAACTGGGCAAGAAATTA 2281
Qy 2276 TCAAGATGTAAAGAGATGAGACGTATGTGATGTGCTGATAGCACTGTGATTCAGTGC 2335
Db 2282 TAAAGAGCTTCAAGAAATATGAAGGCTGAGACGTATGCGCAAGAACTGTGATGCT 2341
Qy 2336 TTCTAAATGGGGTTAAGCACCCCGTTTAACTCTGTATCATTTGATGAGCAATTTGCTGCC 2395
Db 2342 TCTTGAATGTATGTCAAAACCCCGTATGAGACCTGTATTTTGAAGCAAGCTTTTGTGTC 2401
Qy 2396 ATGTAGGAGCGCTGTGCTGACATGATTCGATCGCAACCTTAAAGAAAGTGTATTTGGCG 2455
Db 2402 ATGTAGGACATCTTCAAGCGCTCATACCATCATTAACCTTAAAGGCAAGTGTCTGCG 2461
Qy 2456 GGAAGCCAAAACAATGTGGGCTTTTAAACATGATGTCTGAAAGTACATTTTAAACATG 2515
Db 2462 GGGATCCAAAACAAGTGTGGCTTTTCAATATGATGTCTGAAAGTGCATTTTAAACAG 2521
Qy 2516 ACATATGCACTGAAGTATACATTAAGAACTTCTTGAAGGTGACACAGCTGTAAACG 2575
Db 2522 AGATTTCAAGAGGTCTTCCACAAAAGCAATCTCTGTGTGACATTAATCCGTGACTT 2581

2576 CCATGTCCTCAACGCTCTTCTAACAAGCGAATGAAAGCGTTAACCCATGCTGATA 2635
2582 CGGTGCTCAACCTGTTTATCGACAAAGATGAAAGCGAAGCCGAAAGAGACTA 2641
2656 AATATATATAGATACCAAGGACCAAGAGCCGCAAGAGATATCTGATCTTAACCT 2695
2642 AGATTGAGATTGACCTACTGCGACTACCAAGAGAGAGATCTCATCTCACTT 2701
2696 GTTTCAGAGATGAGTGAACAGCTACAGATTGACTACAAATAACGAAATCATGACTG 2755
2702 GTTTCAGAGATGAGTGAACAGCTACAGATTGACTACAAATAACGAAATCATGACTG 2761
2756 CGGCTGATCGCAAGAGACTACCGGAAAGCGTTATGCTGCTGAGTAAAGTCAACG 2815
2762 CAGCTGCTCAAGAGGCTGACCGGTAAGGCTGATGCTGCTGCTGCTGCTGCTGCTG 2821
2816 AGAATCCACTACTCTGCAAGACTTCTGAGCAGCTGATGCTGCTGCTGCTGCTGCTG 2875
2822 AAAATCCCTGTAACGACCCACCTCAAGACATGTAACGCTCTGACCCGCAAGAGG 2881
2876 AACGATGCTGTAAGAGCGTACGCTGATCCCTGATTAAGACACTTACAGTAAAT 2935
2882 ACCGATGCTGTAAGAGCGTACGCTGATCCCTGATTAAGACACTTACAGTAAAT 2941
2936 ATCCGAGGATTTCAAGGCTTCAATGGAAGAGCTGAGCGGCAAGCAAGCGCATTAATG 2995
2942 ATCCGAGGATTTCAAGGCTTCAATGGAAGAGCTGAGCGGCAAGCAAGCGCATTAATG 3001
2996 CACGCTTCTGTAAGAGCGGCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3055
3002 GGCACATCTGTAAGAGCGGCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3061
3056 GGGGGAAGCTTTAGAGCAAGCTTGGCCACCGCCAAATGCTGCTGCTGCTGCTGCTGCT 3115
3062 GGGGGAAGCTTTAGAGCAAGCTTGGCCACCGCCAAATGCTGCTGCTGCTGCTGCTGCT 3121
3116 GGGGGAAGCTTTAGAGCAAGCTTGGCCACCGCCAAATGCTGCTGCTGCTGCTGCTGCT 3175
3122 GGAACACTGCTGATTAATCTGGAAGAGCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCT 3181
3176 TCTTTTGAACAAGCTTCTGGAAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3235
3182 AACTATGCTGTAAGAGCTTCTGGAAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 3241
3236 TCGCACTTACTTAAGAGAGTACGCACTGCGGATTAATGCTGCTGCTGCTGCTGCTGCTGCT 3295
3242 TTCCGTTTCAATTAAGAGATTAATCACTGCGGATTAATGCTGCTGCTGCTGCTGCTGCT 3301
3296 TTAATAGAGAGTACCAAGAGAGTGTACGCGGATTAATGCTGCTGCTGCTGCTGCTGCTGCT 3355
3302 TGAATTAAGAGAGTGTGTGCGGAGCTGCTGCGGAGTACCGCAACTGCTGCGAGAGTGTG 3361
3356 ACACAGGAGGATGCTGATTAATTAAGAAATTAATACATCAAGGACTACTCTCAACATTA 3415
3362 CCACGGAAGAGTCTATGACATGAACACTGCGGAGCTGCGCAATTAATGATCGGCGCATTA 3421
3416 ATGCTGCTCAATTAATGCGCGGTTGCCCACTGCTGATGCTTACCAACAAGAGAGG 3475
3422 ATCTAGTACTGTGAACAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3481
3476 GTACAACTGATCAACGAGGATTTCTATCTAAGATGAAGGCAAAATCTGTTGCTGATGCTG 3535
3482 CACAGAGTGAATTTCTTCTTCTTCTGCTGAGCAACTGAAGGCGAGAACTGCTTGTGCTGCTG 3541
3536 GCGATCTTATCAAGCTTCCAGGAGAAAGTGAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCT 3595
3542 GGGAGAAAGTGTGCTGCTGCGGCAAAAGAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3601
3596 CCATCAGTGTGATCTGATTTGGGAATACCTAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3655
3602 CTTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3661

3656 TTGTCATGTTAGAGACCCCGTACAGAAACATCACTACCAACAGTGCAGAGATCAAGCTA 3715
3662 TTATTAAGCTGAGAGCTCATATATAATCATCATATATACAGAGTGAAGACAGCGCA 3721
3716 TCACACAGCATGCTTAAGCTGTAAAGCTGTACACACTTGAACACTGCGGAAATGTG 3775
3722 TTAAGCTTAAGTGTGAACAAAGAAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3781
3776 TGGCTATGAGGATGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3835
3782 TCAGCATGCTTATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3841
3836 GCTCATTTAGGTTTACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3895
3842 GCGAGTTCAAGTTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3901
3896 TCTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3955
3902 TGTGTGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3961
3956 TAGTCTGATCAACATCTATCAAGGCTCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4015
3962 CTACCTTACCAACATCTATCAAGGCTCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4021
4016 ACAGAGTATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4075
4022 ATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4081
4076 ATGCAAGGCTCAACAGGCTTCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4135
4082 ACAGCAAGGCTCAACAGGCTTCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4141
4136 CTTTGTGATGAGCAGGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4192
4142 GCTTGTGATTAAGGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4201
4193 TCATCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4252
4202 ATATTTATTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4261
4253 AGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4312
4262 AGTTGCGAGAGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4321
4313 TATCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4372
4322 TAGGATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4381
4373 CATTGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4432
4382 CATTAAGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4441
4433 TGAATTAACATGGAAGACAGAGTAATCGAGGCTTACCGCAAGAAAGCGTGA 4492
4442 GGAACAAGAAATGGAAGATGACTCTCAAGGAAGAGTGTGAGAGAGAGAGTGA 4501
4493 TTCTGGA-----TGATGACAAGCAAGTGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCT 4537
4502 AGAATATGATATCAAGCACTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4561
4538 ACCCAACAGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4597
4562 ATCCGAAGATTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4621
4598 CATACCTGAAGATCAACGATTTCAATGACCGCCCAAGAGCAATGCTGCTGCTGCTGCTGCTGCT 4657
4622 CATATTTGGAAGGAGCAAAATTTCAACGAGGCGCAAGATATGCAAGAAATTAATGCA 4681
4658 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4717
4682 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4741
4718 TGTCCAGATCGCTCCAAATGCGCAAGTGAAGAGTCAAGAGGCTGCTGCTGCTGCTGCTGCTGCT 4777

Dn	4742	TGACGAGATTAGGTCCGAATGCCCGTCGAGAGTGGAAAGCTTCCACACCCTTAGCA	4800
Oy	4778	CAC TTC CAT GGC CTG TTA ATT AAT CAG CTT ATG AGC GCT GAG CGG GTT AAT CAG GTT GCG CTT CTG	4837
Dn	4802	CGCTGCGCTTGCTTGTCATCATCATGCTATGACTTCAGAAAGAGTACAACGCCCTAAAAGGCT	4861
Oy	4838	CGAAGAAAAGAACAGTTCGCCCGGTATGCTATCATCTCTGTTCGCCAAGTACAGGATCACAG	4897
Dn	4862	CACGCTCCAGAACAAATTATCTGTGTCTCATCTCTTCCTTCATTCGCCAAGTATGAAATCACTG	4921
Oy	4898	GCGTGCAAGACCTACAGTGCAGAGCAAAACAGTCTCGTTTTCAGGGCGTGTACACCGGCTG	4957
Dn	4922	GTGTGCAAGACATCCAGTGTCTCCAGCCTATACTGTTCTTCACCGAAGGTGCTGTGCTACA	4981
Oy	4958	TACACCCCGAGAGTAC	4974
Dn	4982	TTTCATCCACGAGATGAC	4998
RESULT 15			
VEU55342		11423 bp RNA linear VRL 15-JAN-2002	
LOCUS		Venezuelan equine encephalitis virus nonstructural polypeptide and structural polypeptide genes, complete cds.	
DEFINITION			
ACCESSION		U55342.2 GI:18152933	
VERSION		1	
KEYWORDS		Venezuelan equine encephalitis virus Venezuelan equine encephalitis virus viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus; VEEV complex. 1 (bases 1 to 11423) Weaver,S.C., Salas,R., Rico-Hesse,R., Ludwig,G.V., Oberste,M.S., Bohehl,J.T. and Tesh,R.B. Re-emergence of epidemic Venezuelan equine encephalomyelitis in South America. VEE Study Group Lancet 348 (9025), 436-440 (1996)	
JOURNAL MEDLINE		96332169	
PUBMED		8709783	
REFERENCE		2 (bases 1 to 11423)	
AUTHORS		Weaver,S.C.	
TITLE		Direct Submission	
JOURNAL		Submitted (18-APR-1996) Scott C. Weaver, Pathology, Univ. TX Med. Branch, 301 University Ave., Galveston, TX 77555-0609, USA	
REFERENCE		3 (bases 1 to 11423)	
AUTHORS		Weaver,S.C.	
TITLE		Direct Submission	
JOURNAL		Submitted (01-MAY-2001) Scott C. Weaver, Pathology, Univ. TX Med. Branch, 301 University Ave., Galveston, TX 77555-0609, USA	
REMARK		Sequence update by submitter	
COMMENT		On Jan 15, 2002 this sequence version replaced gi:235113.	
FEATURES			
SOURCE		location/Qualifiers	
		1..11423	
		/organism="Venezuelan equine encephalitis virus"	
		/mol_type="Genomic RNA"	
		/strain="VI98"	
		/variety="antigenic IC"	
		/db_xref="taxon:11036"	
		25..7506	
		/note="contains low efficiency read through at stop codon: encodes nsP1, nsP2, nsP3 and nsP4 proteins"	
		/codon_start=1	
		/transl_except="(pos:5662..5664,aa:OTHER)"	
		/product="nonstructural polypeptide"	
		/protein_id="AI61965.1"	
		/db_xref="GI:18152933"	
		/translation="MEKVHVDIEEDSPFLALQSFQFEVEAKQVTNDNHANAFSES HLAKLTETEDPSDTIIDIGSAPARARYSGHKYHCIPMKCAEDPRLYKATRLKKK NCKETDELDKKMKEELAAWSDPDLETETNCIDHDSECRVEGOVAOVADVPGLK SLVHOANGKAVIYWGIDTPPMFKNLAGA YPSYTNMADETULTANIGLSSDW ERSRKNSILAKTKLPENNVLFVSGSTITHEKDILRSMLPYFPHLRGNQNTTCRC ETIVSCDGYVVKRIIAISFGILYGKDSGVAAVTHMRBGLCCKVTDTLNERRVSFPVCITY	
CDS			

BASE COUNT 3230 a 2820 c 2900 g 2473 t

ORIGIN

Query Match 42.4% Score 21.9; DB 14; Length 11423;
Best local Similarity 64.5%; Pred. No. 0;
Matches 3223; Conservative 0; Mismatches 1750; Indels 24; Gaps 3;

QY 2 CCTTAACTAATGATCCAAATATGAAAGAATTACGTTGACTTAGATGCTGACAGCC 61
DB 2 CCAAAACCAATTACTTACCCAAAATGAGAAAGTTACAGTTGACATCGAGGAAGACAGCC 61
QY 62 CGTATGTAAGTCGTTACAGGGAGGTTTCCAAATTTGAGATCGAAGCAGGCTCA 121
DB 62 CATTCCTGAGAGCTTTTAAACAGGACCTTCCGACATTTGAGGATGAACCAAGCAGGTCA 121
QY 122 CTGACATGACCATGCGAATGCGAGGGTTTGGCATGTGGCAACAAAGCTCATTTGGA 181
DB 122 CTGATTAATGACCTGATATGCGAGGGGTTTGGCATGTGGCTTCAAAATCGATCGAAA 181
QY 182 GCGAAGTCGACCGGAGCAAGTTATCTTGAATTTGAAGTGCGCCGCTGACAGATGAC 241
DB 182 CGGAGGTGAGCCATCCGACACGATCTTGAATTTGAAGTGCGCCGCCGCCGACAAATGT 241
QY 242 ATTCCAATCACCCTATCATTTGATATGCTCCATATGAGCGGTGAAGACCCGAGACAG 301
DB 242 ATTCTAACATTAAGTATCATTTGATATGCTCCGATGAGATGTGCGGAATCGGACAGAT 301
QY 302 TACAACGCTATGCGAAAGACTTAGAAAGT-----GACATTCGCAAGAAACATATG 355
DB 302 TGTACAAATGATGCACTTAAGCTGAAGAAATTTGCAAGGAAATTACTGACAAAGAAATTGG 361
QY 356 CCTTAAAGCGGAGACCTGCTGAAATGATGTAACACACAGACGAGACGATCCATCTC 415
DB 362 ACAAGAAATTAAGAGAGCTCGCGCCGCTCATGAGGACCCAGACCTGGAATCTGAGATTA 421
QY 416 TGTGATGACACAGACCCACGCTGTAGATCTTTGAAAGTGTAGCAATATACCAATG 475
DB 422 TGTGCTCCACGACATGATGATGCTCCCTTACGAAAGGCAAGTCCGCTTTTACCAGATG 481
QY 476 TGTACGAGCTCATGSCACGACATCAATCCACACAGCGGCTTAAAGAGTTAGGACAA 535
DB 482 TATACGAGTTGACGACCGACGACAAAGTCTCTATCACCAAGCCAAAGGAGTTAAGTGC 541
QY 536 TTTACTGATAGGCTTTTGAACGACCCCTTTATGTACAAAAACATGSCAGTTCTTACC 595
DB 542 CCTACTGATAGGCTTTTGAACGACCCCTTTATGTATTAAGAACTTGGCTGAGATATC 601
QY 596 CTACTTACAAACGAACTGGGCTGACGAGAGATTTGAAAGCAGTAATTTGGCTCG 655
DB 602 CATATACTCTTACCAACTGGGCGGACGAAACCGTGTAAACGCTGTAAACATAGGCTTAT 661
QY 656 GTAACCTGATCTTCAAGSAGCAGGCTTGGAAAACTCTCAATCCTTAGGAAGAAAGAGC 715
DB 662 GCAGCTCGAGCTCATGAGGCGGTCAAGTAGGAGATGCCATTTCTTGAAGAAAGATAT 721
QY 716 TCCAACCTAATTAATGAATCATATTTCTGCTTGTGTTCAACAATCTACACAGAGTAGAT 775
DB 722 TGAACCACTCCAAATATGCTCTATTTCTGTTGGCTTCAACATCTTACACGAAAGAGG 781
QY 776 CACTGTTACGTAAGCTGAGCATCTTCCAAAGTGTTCACCTTGAAGAAAGTCTTATCA 835
DB 782 ACTTACTAGAGAGCTGGCACCTGCGCTGTATTTCACTTACGTGGCAAGCAAAATTACA 841
QY 836 CAGGTAAATGTGGGACCATTTGTCACTGTGAAGGTAGTCAATCAAAAAGTAAAGATCA 895
DB 842 CATGTGGTGTGAGCTTATGTTAGTTTGCAGCGGCTAGCTGTGTTAAAGAAATAGTATCA 901
QY 896 GCCCAGGACTAATCGGTAAAGTTGAGAACTTGGCGTCCACAATGATGCGAGGGTTTCT 955
DB 902 GTCCAGGCTGTATGAGAGGCTTCAAGCTATGCTGCTACGATGACCGCAGAGGATTTCT 961
QY 956 TGAAGTGAAGATCAAGATACGCTGCGGCGAGAGGGTTCCTTTTCTGCTGTGTACGT 1015
DB 962 TGTGCTGAAGATGACAGACATATGAACGCGGAGAGGGTCTCTTTTCCGATGACAGT 1021
QY 1016 ATGTAACGACCACTTTGCGATGAGATGACAGGATTTCTGCACTACATGATGATGAG 1075
DB 1022 ATGTCCAGCTACATTTGTGTGACCAAAATGACTGGCATCTGGCAACAGATGTCAAGTGGG 1081

QY 1076 ATGACGACAAAACTATTGGTTGGGCTCAACCAAGAGATTGCTCAATGTATGAGAGC 1135
DB 1082 ACGAGCGCAAAAACTGCTGGTTGGGCTCAACGAGCGATAGTGTCAACCGTGCACCC 1141
QY 1136 AAAGAAATATTAACAATGACAGAACTATCTATTACCAAGTGTGCGCCAGGCTTTTCCA 1195
DB 1142 AAAGAAACACAAATACCTATGAAGAAATTAATCTTTTGGCCGTAGTGCCAGGCAATTTGCTA 1201
QY 1196 GGTGGGCGGTGAACATGTCGCCGATGAGCGACGAGAAAGAAAGTATAGGGGTGGGGAGC 1255
DB 1202 GTGGGCAAGAAATATTAAGAGATTAAGAAAGTGAAGGCCACTAGGACTACAGATTA 1261
QY 1256 GCACCTTACTATGAGGCTGCTGCTGGGCTTTCAAGACCCAGAAATACATCTCATCTCA 1315
DB 1262 GACAGTTAGTATGAGGGGTGCTGCTGGCTTTTAAAGGCAACAAGTAACTTATTTATA 1321
QY 1316 AGAAGCTGTGTAACGAAACAAATTAAGAAAGTACCTGCCGCTTTGACTCATTTGTGATTC 1375
DB 1322 AGCCGCCAGATACCAAAACCATCATCAAAAGTGAACAGCGATTTCCACTCATTCGTGCTGC 1381
QY 1376 CAGGCTTACAGGCGAGGCTCGATATAGGCTTCCGCCGTAGAGGCTCAAGCTGCTGTTG 1435
DB 1382 CGAGATGAGCATTAACACTGAGATCGGCTGAGAACGAGAAATCAGAAAAATGCTAG 1441
QY 1436 AACCAACTGTCAAAACCCGACACCGCTATTACATATGCGCATGTGAGCATCTGCGTGC 1495
DB 1442 AAGAGCAACAGAGACCGCTACCTCTCATTTACTGCGAGAGACATACAAAGGCTAAGTGC 1501
QY 1496 TACAGCAAGAACTGAAAGATGCTGACGCGAAGAGATCAGAGAACCTTGCACCT 1555
DB 1502 CAGCGGATGAGGCTAAGAGATGCTGTGAACCGAGAGCTCGGCTCCTACATCACACCTT 1561
QY 1556 TGTCTCCCTGAATGAAGAAAGACCCGAGAGGAGAAAGTATGACTCTTATGCAAGAG 1615
DB 1562 TGGAGCTGATTTTGAAGAGCCACTCTGAAAGCGAGTGTGACTGATGTTTACAAAGG 1621
QY 1616 CAGAGCAGTATGAGGTGAGACACACAGAGGACATCAGGATGACAAATTAACCAAGCG 1675
DB 1622 CTGGGGCGGCTCATGTGAGACACTCTGTGCTTATTAAGTTTACAGCTATGCGGCG 1681
QY 1676 AAGAAAGATTGGGCTTTAGCGTATTACTTCAACCCAGGCGGTATTGAATGTGAAAAA 1735
DB 1682 AGGACAAAGATGGGCTTTTACGACATGCTTTTCCACAGGCTGTACTCAAGAGTGAAGAA 1741
QY 1736 TGGGCTGATACACCAATGSCGGAACAAGTACTGTATATGATCACAAGATGAGGAG 1795
DB 1742 TATCTTGTATCACCTCTCTGCTGAACAAGTATGATATTAACAACCTTGGCCGAAAAG 1801
QY 1796 GGAGATACAAAGTGAAGCATATGAGTGAAGAGTCAATGTACAGAAAGGAGCGGCTGC 1855
DB 1802 GCGCTTATGCGGTGGAACCTTACATGAGAAAGTATGCTCCAGAGGACATGCAATAC 1861
QY 1856 CTGTTCAAGCTTCCAGGATTTGAGTGAAGAGCGGTACGATGTTTCAACGAGAGGAGT 1915
DB 1862 CCGTCAAGACTTTCAAGCTCTGAGTGAAGTGCACATCGTGTGAACAACAGAGAGT 1921
QY 1916 TCGTAAACAGTACTCTGACCACTCGCATCAACGAGAGGCGGTAAACACTGACGAAG 1975
DB 1922 TCGTAAACAGGTACTCTGACCATATTTGCCACACATGAGGAGCGCTGAACAGATGAAG 1981
QY 1976 AGTACTATTAAGCTGTAAAGACTCAGGACACAGACTCAGAAATAGTCTTGTGATTTGAG 2035
DB 1982 AATATTAACAACTGTCAAGCTCCAGCGACGACGAGCGAAATACCTGTATACATCGACA 2041
QY 2036 CACGAAAGTGTGTTAAGCGAGAGACGAGGTCCCTTGTGCTTAAACCGGTATCTGTAG 2095
DB 2042 GGAACCAATGCGTCAAGAAAGAAATTAAGTCACTGGGCTTAAAGGCTTAAAGGAGCTGTGG 2101
QY 2096 ATCCACATTTCAAGATTTGCGTACAGAGATCTCAAGACACGACGACGACGACTTACCA 2155
DB 2102 ATCTCTCTTCCATGAATTTGCTTACGAGATCTGAAGAACGTCGCGCTCTTAAAC 2161
QY 2156 AAGTCCAAACCATGAGTATGAGATGCGAGGTTCAAGTAACTGAAATCATCAAAA 2215

Db	2162	AAGTACCAACCATAGGGGTGTATGGCGTCCGGGGTCAAGGCAAGTCTTGACATCTTAAAA	2221
Qy	2216	GCAGTGTGACTAAGAAAAGATCTGGTGTGATGTCGAAGAAAGAAAATCCGCACAGAAATCA	2275
Db	2222	GCGCAGTCAACAAAAGAAAGATCTGGTGGTGAAGCGCAAGAAAGAAAATCTGGCAGAAATTA	2281
Qy	2276	TCAGGGATGTAAAGAGATGAAGATATGTTGCTGCTAGACATCTGCATTCAGTGC	2335
Db	2282	TAAAGGACGTCAAGAAAATGAAAAGGGCGTGAAGCTCAATGCCAGAACCTGTGGACCTCAGTGC	2341
Qy	2336	TTCTAAATAGGGGTTAAGCAACCCCGTTAACCTGTATCATTTAGATGAGGCAATTCGCTGCC	2395
Db	2342	TCCTGAATGATGTCAAACACCCCGTAGAGCCCTGTATATTTAGAGAGACTTTTGCTGTTC	2401
Qy	2396	ATGCAGGAGACCTGCTGTCACATGATTCGCATCTGTCAAACTTAAGAAAGTGTATTTGTGCG	2455
Db	2402	ATGCAGGACCTCTCAGAGCGCTCATATGCACTATAAGCTTAATAAGGCAAGTCTGTGCG	2461
Qy	2456	GGGACCCAAAACAATGCGGCTTTCTTTAACATGATGTGCTTGAAGTACATTTTAAACAATG	2515
Db	2462	GGGATCCAAAACAGTGTGGCTTTTTCATATGATGTGCTTGAAGTACATTTTAAACAACG	2521
Qy	2516	ACATATGACCTGAAGTGTACCATTAAGCATCTCTAGAGAGGTGCACACAGCTGTAAACG	2575
Db	2522	AGATTTGACGACAGGTCTTTCACAAAAGCATCTCTGCTGTGCACTAAATCCGTGACTT	2581
Qy	2576	CCATCGTCTCACGCTCTTCTACGACAAAGCAATGAAGACGGTTAAACCATATGCTGTATA	2635
Db	2582	CGGTGCTTCAACCTTGTTTTACGACAAAAGATGAGAACGACGAAACCCGAAAGAGACTA	2641
Qy	2636	AAATCATATAGATNACACAGGGACACAAAGCCGCAAAAGATGATCTGATTTCAACT	2695
Db	2642	AGATTTGAGATTTGACACTACTGCGAGTACCAAAACGAAACAGACGATCTCATTTCTCACTT	2701
Qy	2696	GTTTCAGAGGATGGGTGAAACAGCTACAGATTTGACTAATAAATACAGAAATTCATGACTG	2755
Db	2702	GTTTCAGAGGATGGGTGAAAGCAGTTGCAAAATAGATTTAACAAGGCAACGAAATTAATGACGG	2761
Qy	2756	CGGCTGATCCGCAAGGACTTAACGCGGAAGGCGTTATGTCTGTCAGGTACAAAGTCAACG	2815
Db	2762	CAGCTGCTCTTCAGAGGCTGACCCGTAAAGGCGGTATATGCCGTTCCGTACAGATGTAATG	2821
Qy	2816	AGATTCACACTCTACCTGCGAGACTTCTAGACAGTGAACGTTGATCTTACACGCACAGAA	2875
Db	2822	AAAAATCCCTGTAGCGCACCCACCTCAGAACATGTGAACGTCCTACTGACCCGCAAGGAG	2881
Qy	2876	AACGATTTGTCTGGAAGACGTAAGTGGTGTATCCTGATTAAGACACTTACAGTAAAT	2935
Db	2882	AACGATTTGTGTGGAAGAACACTATGACCGGTGTATCAATGATTAATAACACTGACGCGCAAGT	2941
Qy	2936	ATCCCGGGGATTTCAACGGCTTCAATTTGACGACCTGGCACGCGGAACACAGCCCAATTATGG	2995
Db	2942	ATCTCTGGGAATTTCACTGCGCACGATGAGAGATGGCAAGACAGCATATGTCATCATGA	3001
Qy	2996	CACGCGTCTTGTATAGCCGCGACAGCAGCTGATGTGTTCACAGATTAAGGTGAACGTCTGCT	3055
Db	3002	GGCACATCTTGGAGAGACCGGACCCCTACGACGTTTCCAAAATTAAGCGAACGTTGT	3061
Qy	3056	GGGCGAAGGCTTTAGAGCCAGTCTTTGGCCACGGCCACATTTGTGCTGACGACAGCAGT	3115
Db	3062	GGGCGAAGGCTTTGTGGTGGCGGTATCGAAGACTGACAGCATATGACATGACCACTGAAACAT	3121
Qy	3116	GGGAGACGTTGACCACTTCAACAGTACGACAGCGCTACCTCAACCTGAATATGGCACTGAAT	3175
Db	3122	GGAAACATGTGATTAATCTTCGAAACGACAAAGCTCACTACGACAAAGATGATTTGAACC	3181
Qy	3176	TCTTTTGACACAGTTCCTTTGGAGTAGACTGGACAGTGGGTTATTTTCCGCTCTCAACG	3235
Db	3182	AACATATGGGTAGGTTCTTTTGGACCTGACCTGGACTCCGGGTGTATTTTCTGCAACCACTG	3241
Qy	3236	TCGCACTTACTTAACGGGATCAGCACTGGGATTAATCGCGACGGGAAGAACATGTATGGGC	3295

Db	3242	TTCCGTTATCATTAAGAAATATACCTGGGATATATCCCGGTGGCTTAACATGTAAGGGT	3301
Qy	3296	TTAAATAGAGAGTAAGCAAAAGAGTTGTCAAGCGGATATCCGTGATCACAAAAGCGTTG	3355
Db	3302	TGAATTAAGAAGTGTGTCCGCAGCTCTCCGCAAGTAACCAACAATCCCTCCGACAGTTG	3361
Qy	3356	ACACAGCAGGGGTACGTGATATAGAAATATACATCAAGGAACTACTCTCCACAATTA	3415
Db	3362	CCACCGGAAGAGTCTTAGCATGACATGACACTGGACAGCTGGCCAAATTAGATCCGGCATAA	3421
Qy	3416	ATGTGTTCCATTAAATCGCCGCTTGCCCACTCGTTGATCGTTGACCAAAAGGACAG	3475
Db	3422	ATCTAGTACCTGTGAACAGAAAGCTGCTCATAGTCTTATGTCTCCACCATATGAAACCC	3481
Qy	3476	GTACAAGTATACAGCGGATTCCTATCTAAGATGAAGGCGAAATCTGTGTGTATCG	3535
Db	3482	CACAGAGTGACTTTCTTCATTGCTGACAAACTGAAGGGCAGAACTGTCTGTGTGCTCG	3541
Qy	3536	GGGATCCATACAGATTTCCAGGGAAGAAATAGAGTCCATGGGTCCATTCGCCCATATA	3595
Db	3542	GGGAGAGTTGTCCGTCCCAAGCAAAAAGGTGCACTGTTGTGACACCACCTGAGGCTA	3601
Qy	3596	CCATCAGGTGTGATCTCGATTTGGGAATACCTACCACTGTGCGAATATATGACATPATCT	3655
Db	3602	CTTTAGAGCTCGGCTGATTTAGGTATCCAGGTGACGTGCCAAATATGACATTTGTAT	3661
Qy	3656	TTGTCAATGTTAGGACCCCGTACAGAACCATCATCAACAACTGCGAGATCAGCTA	3715
Db	3662	TTATTAACGTAGAGACTCCATATATTAATACATCATTTATCAGCAGGTGTAAGAACCCCA	3721
Qy	3716	TCCACACAGATCTCAACGTGTAAAGGTGTGCCACACCTGAAACACTGGGGGAAACATGTG	3775
Db	3722	TTAAGCTTATGATGTGTGACCAAGAAACCTGTGTGCACTTGTGAATCCGGCGGAACCTGCG	3781
Qy	3776	TGGCATATGGGTATGGGCTTGTCTGATGGCGCAACCGAATATATCATCTGCGGTGGCAC	3835
Db	3782	TCAGATATGGTTATGTGTAGCTGTACGTCAACAGGGCAGCGAGAGATATGTGTCTATACGC	3841
Qy	3836	GCTCATTTAGGTTAAACCGGTGTCTGTAGCCTTAAGAAACACTGCCAAATATCTGAGGTTTC	3895
Db	3842	GGCAGTTCAAGTTCTCCCGGGTATGCAAAACGAAATCCTCACTGAAAGACAGAAATGAC	3901
Qy	3886	TCTTCGTGTTTGGGCAAGGACACGGCAACCAACATGACACAGACAGACTCGGTG	3955
Db	3902	TGTTTGTATTTATTTGGGTACATGCGCAAGGCCCTTACGCAAACTCTTACAAAGTTTTCAT	3961
Qy	3956	TAGGTCTTGAACAATCTATCAAGGGTCAACCAAGTACGAGGCGAGGAGAGCTCCACAGT	4015
Db	3962	CTACCTTACCAACATCTATACAGTTTCACAGCTCCACAGAACCGGATGGGCACCTCAT	4021
Qy	4016	ACAGAGTGATCAGAGGTGACATTAGCAAGACGCTGACCAAGCTATCGTTAATGCTGTA	4075
Db	4022	ATCATGTGTGTGCGAGGGGATATTTGCCACAGGCCACGGAAGGATGATCATTAATGCTGCA	4081
Qy	4076	ATAGCAAAAGTCAACCAAGTTCCGGAAGTGTGCGGTGCACTGTACCGAAATATGCGCGCTG	4135
Db	4082	ACAGCAAAAGGCAAACTGTGCGGAAGGGGTGTGCGAGGCGCTGTATTAAGAAATTTCCCGAAA	4141
Qy	4136	CTTTTGAATAGACACCAATAGCTGTGCGGACCGCTAAGCTTGTGAAGCAAGCAACCC---GC	4192
Db	4142	GCTTCGATTTACAGCCGATTCGAAGTGGGAAAAAGCGCATGTGCTAAAGGTATATTAAC	4201
Qy	4193	TCATCATATCATGTGTAGAGCCCAATTTTCTTAAGATCCGGAACCGGAGGCGACCTTA	4252
Db	4202	ATATTATTTCAATCCGTAGGGCCCAACTTCAACAAAGTTTGGGAAGTTGAAGGGGACAAAC	4261
Qy	4253	AGCTCGACGCTGCTACATGACATAGCGTTCATGCTGCAACCTGAAGCGGATTAACAATA	4312
Db	4262	AGTTGGCAGAGCTTATAGATGTCATCGCTAAATTTGTCAACGATTAACAATTAACAGTACG	4321
Qy	4313	TATCAGTACCCCTACTGTCAACCGGCAATATTTCTGGGTGCAAAAGATCGAGTGTGCAAT	4372
Db	4322	TAGCATTTCACTGTTGTGCAACGGGCATCTTTTCCGGGAACAAAGATCGACTTAACCCAT	4381

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 15, 2003, 22:29:23 ; Search time 22769 Seconds
(without alignments)
11651.760 Million cell updates/sec

Title: US-10-023-649-1_COPY_5000_11484

Perfect score: 6485

Sequence: 1 caccgcgcgcacgcacacac.....tatttcttttaaatc 6485

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb ba:*
- 2: gb hcg:*
- 3: gb in:*
- 4: gb om:*
- 5: gb ov:*
- 6: gb pat:*
- 7: gb ph:*
- 8: gb pl:*
- 9: gb pr:*
- 10: gb ro:*
- 11: gb scs:*
- 12: gb sy:*
- 13: gb un:*
- 14: gb vi:*
- 15: em ba:*
- 16: em fun:*
- 17: em hum:*
- 18: em in:*
- 19: em mu:*
- 20: em om:*
- 21: em or:*
- 22: em ov:*
- 23: em pat:*
- 24: em ph:*
- 25: em pl:*
- 26: em ro:*
- 27: em scs:*
- 28: em un:*
- 29: em vi:*
- 30: em hcg_hum:*
- 31: em hcg_inv:*
- 32: em hcg_other:*
- 33: em hcg_mus:*
- 34: em hcg_pln:*
- 35: em hcg_rtd:*
- 36: em hcg_mam:*
- 37: em hcg_vrt:*
- 38: em hcg_hum:*
- 39: em hcg_mus:*
- 40: em hcg_other:*
- 41: em hcg_mus:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6485	100.0	11484	14 AF214040	AF214040 Western e
2	4075.6	62.8	4170	14 EEMSP	U03854 Western equ
3	3850.2	59.4	4047	14 AF228608	AF228608 Western e
4	2705.6	41.7	4692	14 AF339476	AF339476 Highlands
5	2397.6	37.0	11663	6 AR042411	AR042411 Sequence
6	2397.6	37.0	11663	6 AR096557	AR096557 Sequence
7	2397.6	37.0	11663	6 BD071122	BD071122 System fo
8	2396	36.9	11663	6 AR042409	AR042409 Sequence
9	2396	36.9	11663	6 AR096555	AR096555 Sequence
10	2396	36.9	11663	6 BD071120	BD071120 System fo
11	2396	36.9	11663	6 I46902	I46902 Sequence 1
12	2396	36.9	11663	14 ACU38305	U03805 Sindbis-lik
13	2388.4	36.8	11703	6 AR042412	AR042412 Sequence
14	2388.4	36.8	11703	6 AR096558	AR096558 Sequence
15	2388.4	36.8	11703	6 BD071123	BD071123 System fo
16	2386.4	36.8	11717	6 AR042410	AR042410 Sequence
17	2386.4	36.8	11717	6 AR096556	AR096556 Sequence
18	2386.4	36.8	11717	6 BD071121	BD071121 System fo
19	2386.4	36.8	11717	14 ACU38304	U03804 Sindbis-lik
20	2380.4	36.7	11703	14 SINCG	U02363 Sindbis vir
21	2380.4	36.7	11740	6 AR221250	AR221250 Sequence
22	2380.4	36.7	11740	6 AR230677	AR230677 Sequence
23	2380.4	36.7	11740	6 AR234151	AR234151 Sequence
24	2380.4	36.7	11740	6 AR237761	AR237761 Sequence
25	2380.4	36.7	11740	6 BD078845	BD078845 Recombina
26	2380.4	36.7	11740	6 BD085672	BD085672 Recombina
27	2372.2	36.6	11717	14 AF103734	AF103734 Sindbis-1
28	2369.2	36.5	16656	6 AR020969	AR020969 Sequence
29	2369.2	36.5	16656	6 AR043384	AR043384 Sequence
30	2369.2	36.5	16656	6 AR062299	AR062299 Sequence
31	2369.2	36.5	16656	6 AR183758	AR183758 Sequence
32	2365.6	36.5	11708	14 SINCK82	M09205 Ockelbo vir
33	2340.2	36.1	11569	14 AF429428	AF429428 Sindbis v
34	2329.4	35.9	11626	14 AF103728	AF103728 Sindbis v
35	2322.4	35.8	16958	6 AX65480	AX65480 Sequence
36	2227.4	34.3	11678	14 U01034	U01034 Eastern equ
37	2225.2	34.3	11687	6 AX286761	AX286761 Sequence
38	2214.2	34.1	11675	14 EEEVIRNA	X63135 Eastern Equ
39	2198.4	33.9	11684	6 AX286760	AX286760 Sequence
40	2154	33.2	4541	14 AF339474	AF339474 Buggy Cre
41	2150.8	33.2	4525	14 AF339475	AF339475 Fort Morg
42	2122.6	32.7	13905	6 AR137241	AR137241 Sequence
43	2122.6	32.7	13905	6 AR282856	AR282856 Sequence
44	1880	29.0	11824	14 AF162884	AF162884 Aura viru
45	1825.2	28.1	4676	14 AF339477	AF339477 Babanki v

ALIGNMENTS

RESULT 1
AF214040
LOCUS AF214040 11484 bp RNA linear VRL 27-JAN-2000
DEFINITION Western equine encephalomyelitis virus strain 71V-1658
complete cds.
ACCESSION AF214040 AF143811
VERSION AF214040 AF143811
KEYWORDS nonstructural polypeptide and structural polypeptide genes,
SOURCE
ORGANISM Western equine encephalomyelitis virus
Alphavirus; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Neolititzky,D.J., Schmatz,F.L., Parker,M.D., Rayner,G.A.,
REFERENCE
1 (bases 1 to 11484)

Qy 541 CAGAGCGAGTGGACGAGTACATCCGTCAACACTCCAACTGACGGTATGAAAGCGGAG 600
Db 5540 CAGAGCGAGTGGACGAGTACATCCGTCAACACTCCAACTGACGGTATGAAAGCGGAG 5599
Qy 601 CGTATATTTTCTCATCGGAAACAGGCCAAAGTCCACTTTCACAGAAATCAGTACGTCAAT 660
Db 5600 CGTATATTTTCTCATCGGAAACAGGCCCAAGGTCCACTTTCACAGAAATCAGTACGTCAAT 5659
Qy 661 GTAAACTCAAGAACCTTAATTGGATCGGCGCGTCCATGAGAGATTTAAGCCCGCGCC 720
Db 5660 GTAAACTCAAGAACCTTAATTGGATCGGCGCGTCCATGAGAGATTTAAGCCCGCGCC 5719
Qy 721 TCGATCTCGAAAGAGAGAAATGTTCAGAGAAACTGCAATTATGCGCTCTGAGAGAA 780
Db 5720 TCGATCTCGAAAGAGAGAAATGTTCAGAGAAACTGCAATTATGCGCTCTGAGAGAA 5779
Qy 781 ATAGAGCAGGTATCAATCAGGAAAGTGAAGAAATATGAAAGCAATTACAGCGAGCCAC 840
Db 5780 ATAGAGCAGGTATCAATCAGGAAAGTGAAGAAATATGAAAGCAATTACAGCGAGCCAC 5839
Qy 841 TCATTTCTGATTTGGGCACTATCTATCATCATGAGAGTGAATCCTGTCGAGTTCACAG 900
Db 5840 TCATTTCTGATTTGGGCACTATCTATCATCATGAGAGTGAATCCTGTCGAGTTCACAG 5899
Qy 901 TCAATTTCTGTCACCAATCTACTGTCACCGTAAATTACAGGTTTACATCTGACAGAG 960
Db 5900 TCAATTTCTGTCACCAATCTACTGTCACCGTAAATTACAGGTTTACATCTGACAGAG 5959
Qy 961 TCGCGGTTAAACGTGCAACTTATGTTATCAGAGAAATTACCTTACAGTACGCACTTAT 1020
Db 5960 TCGCGGTTAAACGTGCAACTTATGTTATCAGAGAAATTACCTTACAGTACGCACTTAT 6019
Qy 1021 GTATTAACGAGTAACGATGCGTATCTTGAATGCTGACGCGCATCTGCTGTCTAG 1080
Db 6020 GTATTAACGAGTAACGATGCGTATCTTGAATGCTGACGCGCATCTGCTGTCTAG 6079
Qy 1081 ATACAGCACTTTTGTGCGGCTAACTGAGAAAGTCAACCAAGAGCATGCTATTTGC 1140
Db 6080 ATACAGCACTTTTGTGCGGCTAACTGAGAAAGTCAACCAAGAGCATGCTATTTGC 6139
Qy 1141 AGCCAGATTAAGATCAGCCGCTCCATGCGCTTATCAGAAATACATTAACAAATGATTTGG 1200
Db 6140 AGCCAGATTAAGATCAGCCGCTCCATGCGCTTATCAGAAATACATTAACAAATGATTTGG 6199
Qy 1201 CTGCGACTATTAAGAAATGTCACGTTACCCAAATGCGAGAAATTACTGCTTATGATT 1260
Db 6200 CTGCGACTATTAAGAAATGTCACGTTACCCAAATGCGAGAAATTACTGCTTATGATT 6259
Qy 1261 CCGCGGCACTTTAATGTTATGTTTCAAGAAATACGATGCAATGATGATGATGATGAT 1320
Db 6260 CCGCGGCACTTTAATGTTATGTTTCAAGAAATACGATGCAATGATGATGATGATGAT 6319
Qy 1321 CTTTTCGCGATTAACCTTATTCGCGCTTACTACAGAAAGTTCAGCAATATGTCACAAAGC 1380
Db 6320 CTTTTCGCGATTAACCTTATTCGCGCTTACTACAGAAAGTTCAGCAATATGTCACAAAGC 6379
Qy 1381 TGAAGGGCCCAAGACAGATGTTTGGCAATCTATCAATCTTAAACCGTTGCGAG 1440
Db 6380 TGAAGGGCCCAAGACAGATGTTTGGCAATCTATCAATCTTAAACCGTTGCGAG 6439
Qy 1441 AGATACCAATGATCAATTCGTCATGATCTTAAAGAGATGTCAAAGTTACTCCCGGCA 1500
Db 6440 AGATACCAATGATCAATTCGTCATGATCTTAAAGAGATGTCAAAGTTACTCCCGGCA 6499
Qy 1501 CGAAACATACAGAGAGCGGCTTAAAGTTCAGAGTATTCAGGCTGCGATCCCTTGGCTA 1560
Db 6500 CGAAACATACAGAGAGCGGCTTAAAGTTCAGAGTATTCAGGCTGCGATCCCTTGGCTA 6559
Qy 1561 CCGCTTACTTTGCGGAGTTCATCGGGAATTAGTCCGTAGATGAAATGCGGTCTTCTG 1620
Db 6560 CCGCTTACTTTGCGGAGTTCATCGGGAATTAGTCCGTAGATGAAATGCGGTCTTCTG 6619
Qy 1621 CAATATTCATATCTCTCTTCGACATGTACGCGAAGATTTTGAATGATGATTTATGCTGAA 1680

Db 6620 CAATATTCATATCTCTCTTCGACATGTACGCGAAGATTTTGAATGATGATTTATGCTGAA 6679
Qy 1681 ATTTCCACACGGGCAACCAAGTATTTGAACGGAATGCGCTGTTTGAATTAAGCGAAG 1740
Db 6680 ATTTCCACACGGGCAACCAAGTATTTGAACGGAATGCGCTGTTTGAATTAAGCGAAG 6739
Qy 1741 ACGAGCTATGCGCACTTTGCGGCTTGAATGATCTTGAAGACTTGAAGTCCGCAACCGC 1800
Db 6740 ACGAGCTATGCGCACTTTGCGGCTTGAATGATCTTGAAGACTTGAAGTCCGCAACCGC 6799
Qy 1801 TCTTAGATTTGATAGAGCGGCGTTCGCAATATCACACTGTGCACTTACTTACAGGAA 1860
Db 6800 TCTTAGATTTGATAGAGCGGCGTTCGCAATATCACACTGTGCACTTACTTACAGGAA 6859
Qy 1861 CGAGGTTTAAATTTGCTGCTATGATGAATCCGATGATGTTTAAAGCTGTTTCAACA 1920
Db 6860 CGAGGTTTAAATTTGCTGCTATGATGAATCCGATGATGTTTAAAGCTGTTTCAACA 6919
Qy 1921 CACTAGTATATCATGATTTGCTAGCAGATGATCACTAGGAAAGGTTTACAGCTGAGGCT 1980
Db 6920 CACTAGTATATCATGATTTGCTAGCAGATGATCACTAGGAAAGGTTTACAGCTGAGGCT 6979
Qy 1981 GCGCGGCTCTATCGGCGAGCATTAATATGATGATGATGATGATGATGATGATGATGAT 2040
Db 6980 GCGCGGCTCTATCGGCGAGCATTAATATGATGATGATGATGATGATGATGATGATGAT 7039
Qy 2041 CGGAGATGCGGCACTTGGCTGAACATGAGTAAGTAAATTTATGATGATGATGATGAT 2100
Db 7040 CGGAGATGCGGCACTTGGCTGAACATGAGTAAGTAAATTTATGATGATGATGATGAT 7099
Qy 2101 TCAAGCACTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 7100 TCAAGCACTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7159
Qy 2161 GCAAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTGAAGAAACATTTGCAAGCTGATG 2220
Db 7160 GCAAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTGAAGAAACATTTGCAAGCTGATG 7219
Qy 2221 ATACCAGACTGCGACCGCGCGCGGCACTGATGATGAAGCAATGCGATGGAACAGAA 2280
Db 7220 ATACCAGACTGCGACCGCGCGCGGCACTGATGATGAAGCAATGCGATGGAACAGAA 7279
Qy 2281 TTGGAATTAAGGAGAGTGAAGAGGCGGTAAGTTCAGATGATGATGATGATGATGAT 2340
Db 7280 TTGGAATTAAGGAGAGTGAAGAGGCGGTAAGTTCAGATGATGATGATGATGATGAT 7339
Qy 2341 GCTGATCATCAGCTCTCTGTCACAGTGAAGCGGTAAGGCTTAAAGAACTTCAAGACATTA 2400
Db 7340 GCTGATCATCAGCTCTCTGTCACAGTGAAGCGGTAAGGCTTAAAGAACTTCAAGACATTA 7399
Qy 2401 GAGGAGCCCAATCACTCTTACGCGCTGACCTTAAATAGTGAAGTGAAGACGACCT 2460
Db 7400 GAGGAGCCCAATCACTCTTACGCGCTGACCTTAAATAGTGAAGTGAAGACGACCT 7459
Qy 2461 ACCACCGGCGAGATGTTTCCATACCTTCACTTAACTTCAACAGTTTACCTTACAA 2520
Db 7460 ACCACCGGCGAGATGTTTCCATACCTTCACTTAACTTCAACAGTTTACCTTACAA 7519
Qy 2521 TCCGATGCTTACAGGATTCACAAACCTCTTACGCGCGCTGAGAGCGGTTTGGCGCCC 2580
Db 7520 TCCGATGCTTACAGGATTCACAAACCTCTTACGCGCGCTGAGAGCGGTTTGGCGCCC 7579
Qy 2581 GCTGCTGCTCAATTCAGAAATCTTGAAGGATGATGATGATGATGATGATGATGATGAT 2640
Db 7580 GCTGCTGCTCAATTCAGAAATCTTGAAGGATGATGATGATGATGATGATGATGATGAT 7639
Qy 2641 ATACCTTAAATGCGCGCGAGGTCACCGGCAAGAAAGAAAGAGTGTCTTAAAGCCAA 2700
Db 7640 ATACCTTAAATGCGCGCGAGGTCACCGGCAAGAAAGAAAGAGTGTCTTAAAGCCAA 7699
Qy 2701 AACTTACTAGGCTTAAAGAAAGAGAGCAAGCAAGAGAGAGAAAGCAAGCTTAAAGC 2760

Db	7700	ACCTACTCAGCGCTAAAAAGAAAGAACACCAAGCCAAAGAGACGAAACCCGACCTTAAACC	7759
Qy	2761	AGGGAAACGACACACGTATGTGATATGAAAGTTGAGAGTCGGAACAAGACATTTCCGATCATGCT	2820
Db	7760	AGGGAACCGAACACGTATGTATGAAAGTTGAGAGTCGGAACAAGACATTTCCGATCATGCT	7819
Qy	2821	GAACGGCCAACTGMAATGGAATATGCTGCGGTTGTCCGAGAGAAAGGCTGATGAAACCATCTCA	2880
Db	7820	GAACGGCCAACTGMAATGGAATATGCTGCGGTTGTCCGAGAGAAAGGCTGATGAAACCATCTCA	7879
Qy	2881	CGTTGAAGGAATAATTGATTAATGAGCAATTAGCCGCGCTGAAATTGAAGAAGCTAGCAT	2940
Db	7880	CGTTGAAGGAATAATTGATTAATGAGCAATTAGCCGCGCTGAAATTGAAGAAGCTAGCAT	7939
Qy	2941	GTACGACTTGGAGTACGGCGACGTTTCCCAGAAACATGAATCAGACACGCTGCAGTACAC	3000
Db	7940	GTACGACTTGGAGTACGGCGACGTTTCCCAGAAACATGAATCAGACACGCTGCAGTACAC	7999
Qy	3001	CAGGAGCAAAACCAACCGGGCTTCTACACTGGCACACCGGCCCATGCCGATGAGAAATGG	3060
Db	8000	CAGGAGCAAAACCAACCGGGCTTCTACACTGGCACACCGGCCCATGCCGATGAGAAATGG	8059
Qy	3061	GAGATTTTACCGTTACCGAGAGAGTGGGCGGGAAGGCGACAGCGGAACACGATCTTGGCA	3120
Db	8060	GAGATTTTACCGTTACCGAGAGAGTGGGCGGGAAGGCGACAGCGGAACACGATCTTGGCA	8119
Qy	3121	CAACAGAGCGCAGAGTTGTGGCTATTGTTCTAGAGAGTGCAAATGAGGCGACGCGTACGCG	3180
Db	8120	CAACAGAGCGCAGAGTTGTGGCTATTGTTCTAGAGAGTGCAAATGAGGCGACGCGTACGCG	8179
Qy	3181	GCTTTTCAGTGGTCACTTGGAAACGAGAAAGGGGTGACCATTAGGGATACCCCGGAAGTTTC	3240
Db	8180	GCTTTTCAGTGGTCACTTGGAAACGAGAAAGGGGTGACCATTAGGGATACCCCGGAAGTTTC	8239
Qy	3241	TGAACCGGTCGCTCAGTATGAGCGGCTATGGGTCGTTTCSAATGTCACGTTCCCATGCGCA	3300
Db	8240	TGAACCGGTCGCTCAGTATGAGCGGCTATGGGTCGTTTCSAATGTCACGTTCCCATGCGCA	8299
Qy	3301	CAAAACAACCCGTCGTCTATTCACTGACGCGCCAGAAACGAAACATCGACGTCCTCGAAGGAA	3360
Db	8300	CAAAACAACCCGTCGTCTATTCACTGACGCGCGCCAGAAACGAAACATCGACGTCCTCGAAGGAA	8359
Qy	3361	CGTGACACATCCAAATTTGCAACAGCGTCGTGGAGAACGTCCTTGAATATGTCATCACGCGC	3420
Db	8360	CGTGACACATCCAAATTTGCAACAGCGTCGTGGAGAACGTCCTTGAATATGTCATCACGCGC	8419
Qy	3421	GCCCCAAACGAAGCATTTACCGATGACTTCAACCTGACCAAGTCCCTACCTGGGGGTTCTGCC	3480
Db	8420	GCCCCAAACGAAGCATTTACCGATGACTTCAACCTGACCAAGTCCCTACCTGGGGGTTCTGCC	8479
Qy	3481	GTATTTGACGACATCCAAACGCGCGTGTTCAGCCCAATAAATTTGAAACGTGTGGGACGA	3540
Db	8480	GTATTTGACGACATCCAAACGCGCGTGTTCAGCCCAATAAATTTGAAACGTGTGGGACGA	8539
Qy	3541	ATCTGATGATGATGATGATTAGAAATCCAGGTCCTCGGCACAAATTCGCGTCAATCAGGACGG	3600
Db	8540	ATCTGATGATGATGATGATTAGAAATCCAGGTCCTCGGCACAAATTCGCGTCAATCAGGACGG	8599
Qy	3601	CACGCGCGATGTCACCAAAATTCGCTTACATGCTTTTTCGACACGACCAATGACATCAAGGA	3660
Db	8600	CACGCGCGATGTCACCAAAATTCGCTTACATGCTTTTTCGACACGACCAATGACATCAAGGA	8659
Qy	3661	AGACAGTATGAGAGAAATAGCTATCAGCACATCTGSAACCTGCGCGTCTTGGCCACAA	3720
Db	8660	AGACAGTATGAGAGAAATAGCTATCAGCACATCTGSAACCTGCGCGTCTTGGCCACAA	8719
Qy	3721	AGGGTACTTCTGTTAGCTCAATGTCTCCAGGTGACAGTGTAAACCGTCAGTATCAGCAG	3780
Db	8720	AGGGTACTTCTGTTAGCTCAATGTCTCCAGGTGACAGTGTAAACCGTCAGTATCAGCAG	8779
Qy	3781	CGGAGCATCTGAGAAATTCATGCAACCGTGTGAGAGAAAGATTCAGAGGAAAGTTTGTCCGGT	3840
Db	8780	CGGAGCATCTGAGAAATTCATGCAACCGTGTGAGAGAAAGATTCAGAGGAAAGTTTGTCCGGT	8839

QY	3841	AGAGAGTACTTGTTCACCCCTCCATGGAAGCTGTAAAGTCACAGTTTACATCA	3900
DB	8840	AGAGAGACTTGTGTCCACCCCTCATGGAAGCTGTAAAGTCACAGTTTACATCA	8899
QY	3901	CTTGAAGAGACGTCTGCGGGATCTATAACATGCAAGGCCACAGCGGTATTA	3960
DB	8900	CTTGAAGAGACGTCTGCGGGATCTATAACATGCAAGGCCACAGCGGTATTA	8959
QY	3961	GTCTATCTGGAGAAAGCTCAGGCGAAGTACATTAAACCTTCTGGCAAAAGCT	4020
DB	8960	GTCTATCTGGAGAAAGCTCAGGCGAAGTACATTAAACCTTCTGGCAAAAGCT	9019
QY	4021	CACCTGAAGATGAAGTGTGGCGACTACAGACAGGTATGCTGAGCAGCGAAGAAAT	4080
DB	9020	CACCTGAAGATGAAGTGTGGCGACTACAGACAGGTATGCTGAGCAGCGAAGAAAT	9079
QY	4081	GAACGGCTGCACTAAAGCAAAAAGTCACTTGCCTACAGAGCGAACAAAGAAATGGT	4140
DB	9080	GAACGGCTGCACTAAAGCAAAAAGTCACTTGCCTACAGAGCGAACAAAGAAATGGT	9139
QY	4141	CTTGAACCTGCGCGGATCTTATTAAGGACACAGACACTCAGTGAAGGTAAATGCAAT	4200
DB	9140	CTTGAACCTGCGCGGATCTTATTAAGGACACAGACACTCAGTGAAGGTAAATGCAAT	9199
QY	4201	TCCATTCGCGTTGACACCGACAGTCGTGCGCGGTTCGTTAGCTCACAGCGCTACAGTCAC	4260
DB	9200	TCCATTCGCGTTGACACCGACAGTCGTGCGCGGTTCGTTAGCTCACAGCGCTACAGTCAC	9259
QY	4261	GAAGTGGTTCAAAGGCACTACCTTCACCTGACCTGCAATGCGAACCAATTGCTGACAAAC	4320
DB	9260	GAAGTGGTTCAAAGGCACTACCTTCACCTGACCTGCAATGCGAACCAATTGCTGACAAAC	9319
QY	4321	GAGAAATTTGGGGCTGCGAGACAGCAAGCAAGCAATGATTAACAGGCTCTACATCCAG	4380
DB	9320	GAGAAATTTGGGGCTGCGAGACAGCAAGCAAGCAATGATTAACAGGCTCTACATCCAG	9379
QY	4381	GAATTTTTCTGTGGGGCGAGAAAGGGCTGAGTACGATGCGGGTAAACCATGAACCAAGTCAG	4440
DB	9380	GAATTTTTCTGTGGGGCGAGAAAGGGCTGAGTACGATGCGGGTAAACCATGAACCAAGTCAG	9439
QY	4441	AGTCTGGGCCCAAGGAGTCGGACACAGGCGAACCAATGATATGCGCGATAGATCATCAT	4500
DB	9440	AGTCTGGGCCCAAGGAGTCGGACACAGGCGAACCAATGATATGCGCGATAGATCATCAT	9499
QY	4501	CCACTATTATCATGCGGATCCAGTCTACACTGTCTATTGTGCTGTGTGTCGCTCTTTC	4560
DB	9500	CCACTATTATCATGCGGATCCAGTCTACACTGTCTATTGTGCTGTGTGTCGCTCTTTC	9559
QY	4561	TATCTGTGTAGGCACTGTCATCATAGACGTTTGATGCGCAAGCAAGAGAATGCTCT	4620
DB	9560	TATCTGTGTAGGCACTGTCATCATAGACGTTTGATGCGCAAGCAAGAGAATGCTCT	9619
QY	4621	GAGCCCAATGACGCGTTTGACACGAACGAACGATACCAACAGCAATTAAGGTTTTGTGCTG	4680
DB	9620	GAGCCCAATGACGCGTTTGACACGAACGAACGATACCAACAGCAATTAAGGTTTTGTGCTG	9679
QY	4681	CATTGCGCAACCAAGCCTGAACATTGTGAAGAACTTTGAACCATCTGTGTATTAACAA	4740
DB	9680	CATTGCGCAACCAAGCCTGAACATTGTGAAGAACTTTGAACCATCTGTGTATTAACAA	9739
QY	4741	CCAACCGTTTTCTGCGGACAGTTGTGCAATTCCTGTGGACGCGTTGTATTTCTGTTCCG	4800
DB	9740	CCAACCGTTTTCTGCGGACAGTTGTGCAATTCCTGTGGACGCGTTGTATTTCTGTTCCG	9799
QY	4801	CTGCTTTTCAATGCTGCAATGCTTTTTTATTTATTTGTTGACAGGCGTCTGCTGTGGGAAGTAGA	4860
DB	9800	CTGCTTTTCAATGCTGCAATGCTTTTTTATTTATTTGTTGACAGGCGTCTGCTGTGGGAAGTAGA	9859
QY	4861	CGCCTTGAACATGCGACCACTGTGCGCAAAATGTTTCGGGGATCCCGATTAAGCGCTTGTGT	4920
DB	9860	CGCCTTGAACATGCGACCACTGTGCGCAAAATGTTTCGGGGATCCCGATTAAGCGCTTGTGT	9919

QY 4921 CGAACGGGAGTTACGGCCCACTTAACCTGAGATCAAGGTCGTCTCATCGGAATTAC 4980
Db 9920 CGAACGGGAGTTACGGCCCACTTAACCTGAGATCAAGGTCGTCTCATCGGAATTAC 9979
QY 4981 ACCTTCACTAACAAGAGTACGTGACCTGCAAAATTTCCACAGATCTTCTTCCACA 5040
Db 9980 ACCTTCACTAACAAGAGTACGTGACCTGCAAAATTTCCACAGATCTTCTTCCACA 10039
QY 5041 AGTTAATGTCGGGGTCCCTGAGTGAAGGATCTCAAGGGGAGATTACATGCGG 5100
Db 10040 AGTTAATGTCGGGGTCCCTGAGTGAAGGATCTCAAGGGGAGATTACATGCGG 10099
QY 5101 CGTTTGGCGGTGTGATCCCTTTCATGTGGGAGCGCACATCTTCTGTGACATGA 5160
Db 10100 CGTTTGGCGGTGTGATCCCTTTCATGTGGGAGCGCGACATCTTCTGTGACATGA 10159
QY 5161 GAACACAACTGATGAGGGGTACGTGATGCTCTCAAGTCTATATGATCAGC 5220
Db 10160 GAACACAACTGATGAGGGGTACGTGATGCTCTCAAGTCTATATGATCAGC 10219
QY 5221 AGTCGCACTAAAGTTCAACAGTGTCTGAAAGTCGGGCTGATATGATACGGAA 5280
Db 10220 AGTCGCACTAAAGTTCAACAGTGTCTGAAAGTCGGGCTGATATGATACGGAA 10279
QY 5281 CACCAACGCGGACCTGATAGCTTTGTCAATGCGCTACGCCAGGTTCTCAGCGAACCT 5340
Db 10280 CACCAACGCGGACCTGATAGCTTTGTCAATGCGCTACGCCAGGTTCTCAGCGAACCT 10339
QY 5341 GAAGTCTATGACAGGGGCGGATATCAGCGGCTTTTTCACCCCTTTCACCAATAGTGTCTAT 5400
Db 10340 GAAGTCTATGACAGGGGCGGATATCAGCGGCTTTTTCACCCCTTTCACCAATAGTGTCTAT 10399
QY 5401 CAGAAAGGGGCTTTTTCACACTACAGTCTCCCTGATGAGGATGAAACAGAGAGC 5460
Db 10400 CAGAAAGGGGCTTTTTCACACTACAGTCTCCCTGATGAGGATGAAACAGAGAGC 10459
QY 5461 GTTCGGCGATATTCAGGATCTCGCTTATGCTACAGACATAGTACCGGACATGACAT 5520
Db 10460 GTTCGGCGATATTCAGGATCTCGCTTATGCTACAGACATAGTACCGGACATGACAT 10519
QY 5521 ACGGCTGTGAAGCTTTGTGTCAGAAATATCAGTCCCTTACACCCAGAGATATCAGG 5580
Db 10520 ACGGCTGTGAAGCTTTGTGTCAGAAATATCAGTCCCTTACACCCAGAGATATCAGG 10579
QY 5581 GTATGAATGTGAAGAACAACTCAGAGAGAGCCCTGAGAGAACAGACATTTGGATG 5640
Db 10580 GTATGAATGTGAAGAACAACTCAGAGAGAGCCCTGAGAGAACAGACATTTGGATG 10639
QY 5641 TAAATTTGAAGTGAAGGCTCTGCGAGCGCTTAAGTGTCTTACGGGACATCCCTATCTC 5700
Db 10640 TAAATTTGAAGTGAAGGCTCTGCGAGCGCTTAAGTGTCTTACGGGACATCCCTATCTC 10699
QY 5701 GATTGACATCCCTGATGAGCTTTTGTGATCATCAGAAATCACCACAAATTTTGAAGT 5760
Db 10700 GATTGACATCCCTGATGAGCTTTTGTGATCATCAGAAATCACCACAAATTTTGAAGT 10759
QY 5761 TAGTCGACAGTACGAGCTGATTTATTTGCGAGACTTTGGTGTCTTCAATTTACA 5820
Db 10760 TAGTCGACAGTACGAGCTGATTTATTTGCGAGACTTTGGTGTCTTCAATTTACA 10819
QY 5821 GTACAAAGCTGACAGGAGGAGACATTTGTCAAGTCACTCCACAGACAGCTGTTTT 5880
Db 10820 GTACAAAGCTGACAGGAGGAGACATTTGTCAAGTCACTCCACAGACAGCTGTTTT 10879
QY 5881 GAAGGAGCGGACACATGTGATGCTCGGTGAGGACATTAACATTTTACACATC 5940
Db 10880 GAAGGAGCGGACACATGTGATGCTCGGTGAGGACATTAACATTTTACACATC 10939
QY 5941 GAGCCACACAGCAATTTATAGTTTGGCTATGGGGAAGAGTCCACCTGCAATGCTGA 6000
Db 10940 GAGCCACACAGCAATTTATAGTTTGGCTATGGGGAAGAGTCCACCTGCAATGCTGA 10999
QY 6001 ATGTAAACACCGGCGGACACATATATTGAGAACACATTAAGTGAACCAAGATTCCA 6060

Db 11000 ATGTAAACACCGGCGGACACATATTTGAGAAACACATTAAGTGCACCAAGATTCCA 11059
QY 6061 GGCGGAGCTTTCCAAAACATCTTGGAACTGGCTGTGACATGTTTGGGGAGCATATC 6120
Db 11060 GGCGGAGCTTTCCAAAACATCTTGGAACTGGCTGTGACATGTTTGGGGAGCATATC 11119
QY 6121 CCTCATTTGTAGACTTATAGTGTGTGCTGACAGCTCTATGCTTATTAACACAGTAG 6180
Db 11120 CCTCATTTGTAGACTTATAGTGTGTGCTGACAGCTCTATGCTTATTAACACAGTAG 11179
QY 6181 ATGACTAGCGCGGACACATGACATAGCGGTAAACCTGAGTACTTCCAGAAACGCTG 6240
Db 11180 ATGACTAGCGCGGACACATGACATAGCGGTAAACCTGAGTACTTCCAGAAACGCTG 11239
QY 6241 TGCATATGCGGACGCGGCTTGGACATTAACATGATGTATTTCCAGAAAGCAGAGT 6300
Db 11240 TGCATATGCGGACGCGGCTTGGACATTAACATGATGTATTTCCAGAAAGCAGAGT 11299
QY 6301 CATATGCTGTGACAGTGCATTAATGATATCATCTACATTAATTAACACATATAT 6360
Db 11300 CATATGCTGTGACAGTGCATTAATGATATCATCTACATTAATTAACACATATAT 11359
QY 6361 CACTTTATGAGACTCATATGAGTCTTAATATACATACATATTTACTTAAAAAC 6420
Db 11360 CACTTTATGAGACTCATATGAGTCTTAATATACATACATATTTACTTAAAAAC 11419
QY 6421 ACTATACACACTTATTAATCTTTTATATTTTCTTTGTTTATTTGTTTAA 6480
Db 11420 ACTATACACACTTATTAATCTTTTATATTTTCTTTGTTTATTTGTTTAA 11479
QY 6481 ATTTC 6485
Db 11480 ATTTC 11484

RESULT 2
EEMSP
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
FEATURES
source

EEMSP 4170 bp ss-RNA linear VRL 02-AUG-1993
Western equine encephalitis virus RNA encoding structural proteins
at the 3' end of the genome.
J03854
J03854.1 GI:323728
6 kd protein; capsid protein; envelope protein.
Western equine encephalomyelitis virus
Western equine encephalomyelitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; MEV complex.
1 (bases 1 to 4170)
Hahn,C.S., Lustig,S., Strause,E.G. and Strause,J.H.
Western equine encephalitis virus is a recombinant virus
Proc. Natl. Acad. Sci. U.S.A. 85 (16), 5997-6001 (1988)
88320369
3413072
Original source text: Western equine encephalitis virus (strain
BFS1703) RNA, passed in suckling mice and VERO cells, originally
from Cx. tarsalis.
Location/Qualifiers
1..4170
/organism="Western equine encephalomyelitis virus"
/mol_type="genomic RNA"
/db_xref="taxon:11039"
<1..114
/note="nonstructural protein 4"
/codon_start=1
/protein_id="AA42998.1"
/db_xref="GI:323729"
/translation="SREIILAGLIINSLSLTSVNFKSIKNGPITLVG"
159..3869
/note="envelope polyprotein"
/codon_start=1
/protein_id="AA42999.1"
/db_xref="GI:323730"

```
/translation="MPPEPOLNPPVYPTNPMAYRDPNPPRCMRPRPPIAAQIEDL
RSIANTLPKORSNPPGPPPKKKSKSAKPKPTQPKKKKQQAQKTRKPKGRBOM
CMKLESDKTPRIMLNGVNGVACVGRKIMKPLHBEKIDNEQLAAKILKASVYDIE
YGVDPQNMKSDTLQYISDKPFGYNMHHGAVQENGFVYPRGVGGGDSRPLDNK
GRVAIVLGGANEGRITALSVTWNQGVYIKDIPBESBPSLITALCVLSNVPFCD
KPRVCSILABERTLDLLENVDNPNYDILNKLKCSRRPKSITDLPILSPYLDG
CPYCRSHAPCSFPIKINWDESDGSIKIOVSAOFYNOAGTDFKFRYMSFDH
DIKESMDKLIASISGBCORLGHKGYFLAACPGDSVTYISGASBENCTVEKTR
RKPYGREYLFPPPHGLYKCHYVDHLKETSAGIITMHRGPHAYKSYLSEASERYI
KPSGKNTVEKCKGDSYIGVSTRIMNKQCIKASDQTKAVFNSPDILHRT
DHSVQKLIHPEFLTPVCPVPLAHTPTVKMFGILHLATPTLLITRKGLARD
ATAEMITGTSRNFVSVEBGLLEVYMGHNEPVRVWAQESAPDPGMGHEI11HYHSH
PWTYIVILGCVALLIVGTASASACIKARBDCLTPALAPNAVFPALVALACIRPT
TOAVSGYEMKNNGSRPLQETAPFGCKIEVPLASACAGH1RISIDIPDAFVRS
ESPTLIEVSCVADCIYSADPGSLTIQYKADRBGHCPVSHSHSTAVLKXATHTVAV
GSTLHFSTSPQANFIVSLCGKKTTCNAEKKPADH1IGEPHAKVDQFQAAVSKTSM
NMLIALFGASSLIVGLIVLVCSSMLINTR"
mat_peptide
/product="capsid protein"
159..935
mat_peptide
/product="envelope protein E3"
936..1115
mat_peptide
/product="envelope protein E3"
1116..2384
/product="envelope glycoprotein E2"
2385..2549
mat_peptide
/product="6 kd protein"
2550..3866
mat_peptide
/product="envelope glycoprotein E1"
3867..4115
BASE COUNT 1157 a 1080 c 992 g 941 t
ORIGIN
```

```
Query Match 62.8%; Score 4075.6; DB 14; Length 4170;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 411; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
QY 2316 TCAGATATCGATGATCTGCGAGGCTGATCATCATCGTCTGTGTCAGGTTAGCCGAA 2375
Db 1 TCAGATATCGATGATCTGCGAGGCTGATCATCATCGTCTGTGTCAGGTTAGCCGAA 60
QY 2376 AGCGTTAAGACTTCAAGAGATTAAGAGAGGCCCAATCACCCTCTAGGCTGACTTAA 2435
Db 61 AGCGTTAAGACTTCAAGAGATTAAGAGAGGCCCAATCACCCTCTAGGCTGACTTAA 120
QY 2436 TAGGTGACGTAGTACACGACGACCTACCCGCGAGAAATGTTTCATATACCTCAGCTGA 2495
Db 121 TAGGTGACGTAGTACACGACGACCTACCCGCGAGAAATGTTTCATATACCTCAGCTGA 180
QY 2496 ACTTTCACCACTTACCTTCAAAATCCGATGCTTACCGAGATCCAAACCTCTCTAGGC 2555
Db 181 ACTTTCACCACTTACCTTCAAAATCCGATGCTTACCGAGATCCAAACCTCTCTAGGC 240
QY 2556 GCCCGTGAAGGCGGTTGCGCCCGCGTGGCTGCTCAATGGAAGATTTTGAAGAGTGA 2615
Db 241 GCCCGTGAAGGCGGTTGCGCCCGCGTGGCTGCTCAATGGAAGATTTTGAAGAGTGA 300
QY 2616 TAGTCACTTGAAGTTCCTTCAAAACGATCACTTAATCCGCGCGAGGTCAACGCGCAAGA 2675
Db 301 TAGTCACTTGAAGTTCCTTCAAAACGATCACTTAATCCGCGCGAGGTCAACGCGCAAGA 360
QY 2676 AGAAGAGAGTGTCTCTTAAGCCAAACCTACTCAGCTTAAAGAAAGAGAGAGAGCA 2735
Db 361 AGAAGAGAGTGTCTCTTAAGCCAAACCTACTCAGCTTAAAGAAAGAGAGAGAGCA 420
QY 2736 AGAGAGAGAGTGTCTCTTAAGCCAAACCTACTCAGCTTAAAGAAAGAGAGAGAGCA 2795
Db 421 AGAGAGAGAGTGTCTCTTAAGCCAAACCTACTCAGCTTAAAGAAAGAGAGAGAGCA 480
QY 2796 CGAGCAAGACATTTCCGATCATGTGAACGCGCAAGTGAATGATATCTGCGTGTGCG 2855
```

```
Db 481 CGAGCAAGACATTTCCGATCATGTGAACGCGCAAGTGAATGATATCTGCGTGTGCG 540
QY 2856 GAGAGAGAGTGTCTCTTAAGCCAAACCTACTCAGCTTAAAGAAAGAGAGAGAGCA 2915
Db 541 GAGAGAGAGTGTCTCTTAAGCCAAACCTACTCAGCTTAAAGAAAGAGAGAGAGCA 600
QY 2916 CCGTGAAGTGAAGAGAGTGTCTCTTAAGCCAAACCTACTCAGCTTAAAGAGAGAGCA 2975
Db 601 CCGTGAAGTGAAGAGAGTGTCTCTTAAGCCAAACCTACTCAGCTTAAAGAGAGAGCA 660
QY 2976 TGAATCAGACAGCTGTCACTACCAAGCAACCAACCGGCTTCACTGAGCAAC 3035
Db 661 TGAATCAGACAGCTGTCACTACCAAGCAACCAACCGGCTTCACTGAGCAAC 720
QY 3036 ACGCGGAGTCCAGTATGAAGATTTGAGGAGATTTGAGGAGAGAGAGAGAGAGAGAG 3095
Db 721 ACGCGGAGTCCAGTATGAAGATTTGAGGAGATTTGAGGAGAGAGAGAGAGAGAGAG 780
QY 3096 GCGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3155
Db 781 GCGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 3156 GTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3215
Db 841 GTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 3216 CCAATTAGAGATACCCCGGAGAGTGTGAAACCGGTGTCACTAGTAAAGAGAGAGAGAG 3275
Db 901 CCAATTAGAGATACCCCGGAGAGTGTGAAACCGGTGTCACTAGTAAAGAGAGAGAGAG 960
QY 3276 TTTGCAATGATCAGGTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3335
Db 961 TTTGCAATGATCAGGTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 3336 GAACACTGACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3395
Db 1021 GAACACTGACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 3396 ACGCTTGAAGATGTCATCAGCGCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3455
Db 1081 ACGCTTGAAGATGTCATCAGCGCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 3456 CCAATCCCTACTGAGGAGTGTGCGCCGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 3515
Db 1141 CCAATCCCTACTGAGGAGTGTGCGCCGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 3516 TAAAAATTTGAAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3575
Db 1201 TAAAAATTTGAAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 3576 CACAATTCGCTCAATCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3635
Db 1261 CACAATTCGCTCAATCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 3636 TCGACACAGACATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3695
Db 1321 TCGACACAGACATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 3696 GACCTGCGGTGCTTGGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3755
Db 1381 GACCTGCGGTGCTTGGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 3756 ACAATGTAACCGTACAGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3815
Db 1441 ACAATGTAACCGTACAGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 3816 AGATCAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3875
Db 1501 AGATCAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 3876 TGTGTAAGTGCAGGTTTACGATCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3935
Db 1561 TGTGTAAGTGCAGGTTTACGATCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
```


QY 3936 ACAGGCCAGGCCCAACGCGTAAAGTCTTATCTGGAGGAAGCGTCAGGCGAAGTGTACA 3995
Db 1621 ACAGGCCAGGCCCAACGCGTAAAGTCTTATCTGGAGGAAGCGTCAGGCGAAGTGTACA 1680
QY 3996 TTTAAACCACTTCTTGGCAAGAACTGACCTTACGAATGTAAAGTGGCCACTACAGCACAG 4055
Db 1681 TTTAAACCACTTCTTGGCAAGAACTGACCTTACGAATGTAAAGTGGCCACTACAGCACAG 1740
QY 4056 GTATGATGAGCAGCGGAACGAAGATGAACGGCTGCACTTAAACCAAAACAGTGCATTTGGCT 4115
Db 1741 GTATGATGAGCAGCGGAACGAAGATGAACGGCTGCACTTAAACCAAAACAGTGCATTTGGCT 1800
QY 4116 ACAAGAGCGACCAAAACGAATGGGTCTTCACTCGCCGGATCTTATTAAGGACACAGAC 4175
Db 1801 ACAAGAGCGACCAAAACGAATGGGTCTTCACTCGCCGGATCTTATTAAGGACACAGAC 1860
QY 4176 ACTCAGTCGAAGTAAATTTGCACTTCCGCTTGAACACCGACAGTCTGCCGGTTC 4235
Db 1861 ACTCAGTCGAAGTAAATTTGCACTTCCGCTTGAACACCGACAGTCTGCCGGTTC 1920
QY 4236 CGTTAGCTCACAACGCTTACAGTCAAGTGGTTTCAAGGCACTACCTTCCACTGACTG 4295
Db 1921 CGTTAGCTCACAACGCTTACAGTCAAGTGGTTTCAAGGCACTACCTTCCACTGACTG 1980
QY 4296 CAATGCGACCAACATTTGCTGACAAAGAGAAATTTGGGGCTGCGACAGACGCAACAG 4355
Db 1981 CAATGCGACCAACATTTGCTGACAAAGAGAAATTTGGGGCTGCGACAGACGCAACAG 2040
QY 4356 AATGATTAACAGAGTCTTACATTCAGAGAAATTTTCTGTGGGGGAGAGAGGGCTGAGTACG 4415
Db 2041 AATGATTAACAGAGTCTTACATTCAGAGAAATTTTCTGTGGGGGAGAGAGGGCTGAGTACG 2100
QY 4416 TATGGGGTAAACATTAACAGTCAAGTCTGGGCCCAAGATCGGCAACAGGCGCAAC 4475
Db 2101 TATGGGGTAAACATTAACAGTCAAGTCTGGGCCCAAGATCGGCAACAGGCGCAAC 2160
QY 4476 ATGATGCGCCGATGAGATCATCATCTCACTATTAATCATCGGCACTCAGTCTCACTGTCA 4535
Db 2161 ATGATGCGCCGATGAGATCATCATCTCACTATTAATCATCGGCACTCAGTCTCACTGTCA 2220
QY 4536 TTGTGCTGTGTGTGTGTGTCTTGTCTATCTGTGAGGCACTGCATCATCAGACGCTTCA 4595
Db 2221 TTGTGCTGTGTGTGTGTGTCTTGTCTATCTGTGAGGCACTGCATCATCAGACGCTTCA 2280
QY 4596 TCGCCAAAGCAAGAAAGATGCTGAGCCCTGAGCCCTAGCCCTGCAACCGAAGCGTAC 4655
Db 2281 TCGCCAAAGCAAGAAAGATGCTGAGCCCTGAGCCCTAGCCCTGCAACCGAAGCGTAC 2340
QY 4656 CCACAGCATTAAGCGGTTTGTGCTGATTCGCGCAACCAAGCGTGAACATTTGGAGAA 4715
Db 2341 CCACAGCATTAAGCGGTTTGTGCTGATTCGCGCAACCAAGCGTGAACATTTGGAGAA 2400
QY 4716 CTTTGAACCATCTGTGTTTAAACAACAACGTTTCTGTGGCAACAGTTGTGATCTCTC 4775
Db 2401 CTTTGAACCATCTGTGTTTAAACAACAACGTTTCTGTGGCAACAGTTGTGATCTCTC 2460
QY 4776 TGGCAGCGCTTGTATTTCTGTTCGGCTGCTTTCATGCTGCACTGCTTTTATTTGTTG 4835
Db 2461 TGGCAGCGCTTGTATTTCTGTTCGGCTGCTTTCATGCTGCACTGCTTTTATTTGTTG 2520
QY 4836 CAGGCGTCTGCTGGGGAAGGTAGAGCGCTTGCAATGCGACCACTGCGCAAAATGTTTC 4895
Db 2521 CAGGCGTCTGCTGGGGAAGGTAGAGCGCTTGCAATGCGACCACTGCGCAAAATGTTTC 2580
QY 4896 CGGGATCCCGTATTAAGCGTTGTGCAACGCGAGGTTTACGCGCACTTAACTTGAGAG 4955
Db 2581 CGGGATCCCGTATTAAGCGTTGTGCAACGCGAGGTTTACGCGCACTTAACTTGAGAG 2640
QY 4956 TCAAGGTGTCTCATCGGAATTAACAACCTTCACTTAACAAGAGTACGTGACCTGCAAT 5015
Db 2641 TCAAGGTGTCTCATCGGAATTAACAACCTTCACTTAACAAGAGTACGTGACCTGCAAT 2700

QY 5016 TCACACAGTCAATTCCTTCAACCAAGTTAAATGTGCGGGTCCCTCGAGTCAAGGAT 5075
Db 2701 TCACACAGTCAATTCCTTCAACCAAGTTAAATGTGCGGGTCCCTCGAGTCAAGGAT 2760
QY 5076 CCTCAAAGCGGATTTACATATGCGCGCTTTTGGCGGTGTGTAACCTTTTATGTGGGAG 5135
Db 2761 CCTCAAAGCGGATTTACATATGCGCGCTTTTGGCGGTGTGTAACCTTTTATGTGGGAG 2820
QY 5136 GCGCAATGTCTGTGTGACAGTGAACACACAACCTAAGTGAAGCGGTACGTCGAGTTCG 5195
Db 2821 GCGCAATGTCTGTGTGACAGTGAACACACAACCTAAGTGAAGCGGTACGTCGAGTTCG 2880
QY 5196 CTCGACAGTGCATTAAGATACAGCAGTCGCACTTAAAGTTTCAACACAGCTGCTGAAAG 5255
Db 2881 CTCGACAGTGCATTAAGATACAGCAGTCGCACTTAAAGTTTCAACACAGCTGCTGAAAG 2940
QY 5256 TCGGCTGCTGTATAGTATACGGCAACACACCGCGCACCTGGAATAGTTTGTCAATGGCG 5315
Db 2941 TCGGCTGCTGTATAGTATACGGCAACACACCGCGCACCTGGAATAGTTTGTCAATGGCG 3000
QY 5316 TCAAGCGAGTTCCTCAACGGGACCTGAAGGTCAATAGAGGCGCGATATCAAGCGCTTTT 5375
Db 3001 TCAAGCGAGTTCCTCAACGGGACCTGAAGGTCAATAGAGGCGCGATATCAAGCGCTTTT 3060
QY 5376 CACCTTTGACCATTAAGGTGCTCATCAGAAAGGGGCTGTTTCAACTACGACTTCCCTG 5435
Db 3061 CACCTTTGACCATTAAGGTGCTCATCAGAAAGGGGCTGTTTCAACTACGACTTCCCTG 3120
QY 5436 AGTATGAGCTATGAAACAGAGAGGTTGCGCGATATTAAGCATCTCGCTTGAATGCTA 5495
Db 3121 AGTATGAGCTATGAAACAGAGAGGTTGCGCGATATTAAGCATCTCGCTTGAATGCTA 3180
QY 5496 CAGACATGTATGACCCCGCATGCACTACGAGTGTGAAAGCTTCTGTCAAGAACTCCACG 5555
Db 3181 CAGACATGTATGACCCCGCATGCACTACGAGTGTGAAAGCTTCTGTCAAGAACTCCACG 3240
QY 5556 TCCCTTACACCAAGCAGATATCAGGATGTAAGTGAAGAAACAACCTCAGAGACGCC 5615
Db 3241 TCCCTTACACCAAGCAGATATCAGGATGTAAGTGAAGAAACAACCTCAGAGACGCC 3300
QY 5616 TGCAGAAACAGCAACATTTGATGTAAATTTGAAGTGAAGCTTCTGAGACGCTTAACT 5675
Db 3301 TGCAGAAACAGCAACATTTGATGTAAATTTGAAGTGAAGCTTCTGAGACGCTTAACT 3360
QY 5676 GTGCTTACGGGCAATCCCTATCTCGATTAATCCCTGAGACGACCTTTGTGAGATCAT 5735
Db 3361 GTGCTTACGGGCAATCCCTATCTCGATTAATCCCTGAGACGACCTTTGTGAGATCAT 3420
QY 5736 CAGATCAACAACATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5795
Db 3421 CAGATCAACAACATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3480
QY 5796 ACTTTGTGTCTCTTAAACATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5855
Db 3481 ACTTTGTGTGTCTCTTAAACATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3540
QY 5856 ACTCCCACTCAACAGACGCTGTTTGAAGAAAGGACCAACATGTGACTGCGGTAGGCA 5915
Db 3541 ACTCCCACTCAACAGACGCTGTTTGAAGAAAGGACCAACATGTGACTGCGGTAGGCA 3600
QY 5916 GCATTAACATTAATTTAGCATGAGCCCAAGCAAGCAATTTTATAGTTTGCATAGCG 5975
Db 3601 GCATTAACATTAATTTAGCATGAGCCCAAGCAAGCAATTTTATAGTTTGCATAGCG 3660
QY 5976 GCAAGAACTCACTGCAATGCTGAATGTAAACACCGCGCAACCACTTAATTTGAGAGAC 6035
Db 3661 GCAAGAACTCACTGCAATGCTGAATGTAAACACCGCGCAACCACTTAATTTGAGAGAC 3720
QY 6036 CACATAAAGTGAACCAAGAAATTTCAAGGCGGCAAGTTTCCAAAACATTTGGAAATGGCTGC 6095
Db 3721 CACATAAAGTGAACCAAGAAATTTCAAGGCGGCAAGTTTCCAAAACATTTGGAAATGGCTGC 3780
QY 6096 TTGCACTGTTTGGGGAGCATCATCCCTCATTTGTGTGAGACTATATGTGTGCTGCA 6155

```
Db      3781 TTGACGTTGGGGGAGCATCATCCCTCATGTTGTATAGACCTATATGTGTGCTGCA 3840
Qy      6156 GCTTATAGCTTATTAACAACACAGTATGATGACGACGGGAGACATGACATAGCGGTAAAC 6215
Db      3841 GCTTATAGCTTATTAACAACAGTATGATGACGACGGGAGACATGACATAGCGGTAAAC 3900
Qy      6216 TCGATGACTTCCGAGGAGCGTGTGATTAATGCAACGGCGGCTTGACACTAAACTC 6275
Db      3901 TCGATGACTTCCGAGGAGCGTGTGATTAATGCAACGGCGGCTTGACACTAAACTC 3960
Qy      6276 GATGATATTTCCGAGGAGACACAGTATGATGACGACGGGAGACATGACATAGCGGTAAATC 6335
Db      3961 GATGATATTTCCGAGGAGACACAGTATGATGACGACGGGAGACATGACATAGCGGTAAATC 4020
Qy      6336 ACACCTACATATTAACAACACATATATCACTTTATGAGACTGACATAGGGTCTTAATATA 6395
Db      4021 ACACCTACATATTAACAACACATATATCACTTTATGAGACTGACATAGGGTCTTAATATA 4080
Qy      6396 CACTACACATATTTACTTAAACACATATACACTTAAATCTTATTAATTTT 6455
Db      4081 CACTACACATATTTACTTAAACACATATACACTTAAATCTTATTAATTTT 4140
Qy      6456 CTTTGTGTTTATTTGTTTAAATTTT 6485
Db      4141 CTTTGTGTTTATTTGTTTAAATTTT 4170

RESULT 3
AF229608 4047 bp RNA linear VRL 27-MAR-2000
LOCUS Western equine encephalomyelitis virus structural polyprotein gene,
DEFINITION complete cds.
ACCESSION AF229608
VERSION AF229608.1 GI:7330240
KEYWORDS Western equine encephalomyelitis virus
SOURCE Western equine encephalomyelitis virus
ORGANISM Alphavirus; WEV complex.
REFERENCE 1 (bases 40 to 816)
AUTHORS Uryvaev,L.V., Lebedev,A.Iu., Sokolova,T.M. and Inferov,V.P.
TITLE Primary structure of the nucleocapsid gene C and the protein coded
by it from the Western equine encephalomyelitis virus
JOURNAL Dokl. Akad. Nauk. 344 (3), 397-401 (1995)
MEDLINE 96097271
PUBMED 8535272
REFERENCE 2 (bases 40 to 816)
AUTHORS Uryvaev,L.V. and Lebedev,A.Iu.
TITLE Comparative analysis of primary structure of nucleocapsid protein
from Western equine encephalomyelitis virus and other alphaviruses
JOURNAL Vopr. Virusol. 41 (6), 252-259 (1996)
MEDLINE 97130212
PUBMED 899681
REFERENCE 3 (bases 997 to 2265)
AUTHORS Uryvaev,L.V., Lebedev,A.Iu., Sokolova,T.M. and Inferov,V.P.
TITLE Capsid glycoprotein of B2 Encephalitis virus, Western Equine:
JOURNAL Dokl. Akad. Nauk. 357 (1), 134-139 (1997)
MEDLINE 9471275
PUBMED
REFERENCE 4 (bases 1 to 4047)
AUTHORS Uryvaev,L.V.
TITLE Primary structure of WEV 26S RNA
JOURNAL Unpublished
REFERENCE 5 (bases 1 to 4047)
AUTHORS Uryvaev,L.V.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2000) Molecular Biology, D. I. Ivanovsky
Institute of Virology, 16 Gamaleya str., Moscow, Russia
FEATURES
source
1..4047
/organism="Western equine encephalomyelitis virus"
```

```

CDS
/mol_type="genomic RNA"
/strain="McMillan"
/db_xref="taxon:11039"
40..3747
/codon_start=1
/product="structural polyprotein"
/protein_id="AF60166.1"
/db_xref="GI:7330241"
/translation="MPPYPOLNPPVYPTNPMAYRDPNPPRCMRPRPPLAQTEDL
RRTSIAUTLTKQRPANPPPCPPPKKSSAPPRKTKQKKQAKTKRKPKPKRQM
CMLESDDKTPPLMLNGVNGVACVVGRLMKPLHVGSKLDNEQLAKLKASMYDLS
YGVVPLNMKSDTLQYTSDKPPGYNMHGVQYENRFTVPRGVGKGSGRPLDR
GRVAVLVGANGRTALSVVMQKGVITKDPGSEBSWLVTLVSNVTFPDC
KPPVCSLAPERTDLVLENDVNDPNTLLENVLKCPSPKSIITDDFLTSPYLG
CPNCRSAPCPSPKILENWDSDGSIRVOVAOYGAGTADTKPYMYSVDH
DIEDSMEKAIATSGPCRLGKGYFLAQCRRGSVYTSISGSENSCTYEFKR
RKFGVEEYLFPVPHKRYVCHVYDHLKESAGYITMHRSPHAYSVYLEASGEVIT
KPPSGKNVLYECKGDSYSGIYSTPKMNGCTAKOCIAKSDQTKWVNSPDLIRAT
DHSVQKLNHPRLIPVCPVPLAPLVTKMFKGITLHLATRPILLTKRGLRAD
ATSEWITGTTSRNFSVGRGLEEYVMGNHBEVRVAOESAPSDHGMPEHIIHYHRH
PVYTVLVCGVALAILVGRASSACITKARDCLTPYALANATVPTALALCIRPT
MATPFGTTLNHLFNNQPLMAQLCIPALAILFRCFSCEMFLVAGCLKVDNF
EHAFTVPNVGIPYKALVERAGYAPLNLEIKVSSLEUTSTNKEYTCFKHYIPSO
VKCCGSLIECKASKADYTCRVFGVYEFPMWGAQCCDSSENTQLSEAYFEAPCTID
HAVALRVHTALRVGLRIYVGNNTAYLDTFVNGVTPGSSRDLVIAGPISAESPDH
KVIVIRGLVYNDPPEYGAAMPKAPGDIOASSLDATDVARIDRIILKPEVKIHWY
TOAVSGEYEMKNNSGRPLEETAPGCKIEVEPRASNCAYGHIPISIDPDAAVFRS
ESPFTLIEVSTADCIYSADPGSLTQYKADREGCHCPYHSHSTTAYLEATHTYTAG
SITLHPSTSSPQANFVSLCGKTKTCAECKPPADHIIGBPHVDOEFOAAVSKTSIN
MLTLFPGASSLIVGLVIVVSSMLINTR"
40..816
/product="C-protein"
/note="nucleocapsid protein"
/citation= [1]
/citation= [2]
817..996
/product="E3-polypeptide"
997..2265
/product="E2-protein"
/note="envelope glycoprotein"
/citation= [3]
2266..2460
/product="6K-polypeptide"
2461..3744
/product="E1-protein"
/note="envelope glycoprotein"
3748..4047
BABS COUNT 1125 a 1041 c 962 g 919 t
ORIGIN
Query Match 59.4%; Score 3850.2; DB 14; Length 4047;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 3939; Conservative 0; Mismatches 108; Indels 4; Gaps 2;
Qy 2435 ATAGTGACCTAGTAGACAGCACCTACCCAGCGGAGAAATGTTTCATACCTCACTG 2494
Db 1 ATAGCTACCTAGTAGACAGCACCTACCCAGCGGAGAAATGTTTCATACCTCACTG 60
Qy 2495 AACTTTCACCACTTACCTTACCAATCGATGCTTACCGAGATCCAAACCTCTCTAGG 2554
Db 61 AACTTCCCACTTACCTTACCAATCCGATGCTTACCGAGATCCAAACCTCTCTAGG 120
Qy 2555 CGCCGCTGAGAGCGCTTGGCCCCCGCGTGGCTCAATTCGAAGTCTTAGAGGTCG 2614
Db 121 TGGCGCTGAGAGCGCTTGGCCCCCGCGTGGCTCAATTCGAAGTCTTAGAGGTCG 180
Qy 2615 ATAGTCACTTGACTTCAACAACAGATCACTTAATCCGCGCAGGTCCACGCCCAAG 2674
Db 181 ATAGCACTTAACTTTCAACAACAGACACTTAATCCGCGCAGGTCCACGCCCAAG 240
Qy 2675 AAGAAGAAAGTCTCTTAAGCAAAACCTACTGAGCTTAAAAAAGAAAGACCAAGCC 2734
Db 241 AAGAAGAAAGTCTCTCTTAAGCAAAACCTACTGAGCTTAAAAAAGAAAGACCAAGCC 300
```

```
QY 2735 AAGAGACGAAACGCAACCTTAAACAGAGAAACGACACGTATGTATGAAAGTTGAG 2794
Db 301 AAGAGACGAAACGCAACCTTAAACAGAGAAACGACACGTATGTATGAAAGTTGAG 360
QY 2795 TCGGCAAGACATTTCCGATCATGCTGAAACGCGCAAGTGAATGATATGCTGCTGTC 2854
Db 361 TCGGCAAGACATTTCCGATCATGCTGAAACGCGCAAGTGAATGATATGCTGCTGTC 420
QY 2855 GGAGGAGGCTGATGAAACCACTCCAGCTGAAAGGAAAAATTGATATGAGCAATTAAG 2914
Db 421 GGAGGAGGCTGATGAAACCACTCCAGCTGAAAGGAAAAATTGATATGAGCAATTAAG 480
QY 2915 GCCGTAATTTGAAGAGGCTAGCATGTATGACGCTTGAGTACGAGGCTTCCAGAAC 2974
Db 481 GCCGTAATTTGAAGAGGCTAGCATGTATGACGCTTGAGTACGAGGCTTCCAGAAC 540
QY 2975 ATGAATCAGACACGCTGCAATGACACGAGCAAAACAACCGGCTTCAACCTGAC 3034
Db 541 ATGAATCAGACACGCTGCAATGACACGAGCAAAACAACCGGCTTCAACCTGAC 600
QY 3035 CACGCGAGCTCCAGTATGAGATGGGATTTACCGTACCGAGAGAGTGGCGGAAA 3094
Db 601 CACGCGAGCTCCAGTATGAGATGGGATTTACCGTACCGAGAGAGTGGCGGAAA 660
QY 3095 GCGGACAGCGGAAAGCCGATCCTGGAACAAGAGGAGAGTGGCTATTGTTTGAAGA 3154
Db 661 GCGGACAGCGGAAAGCCGATCCTGGAACAAGAGGAGAGTGGCTATTGTTTGAAGA 720
QY 3155 GGTGCAAAATGAGGAGCCGCTGACGGGCTTTCACTGCTGACCTTGAACCAAGAGGCTG 3214
Db 721 GGTGCAAAATGAGGAGCCGCTGACGGGCTTTCACTGCTGACCTTGAACCAAGAGGCTG 780
QY 3215 ACCATTAGGATATCCCCCGAAGTTCTGAACCGTGTGTACTAGTATGACGCTAGCTG 3274
Db 781 ACCATTAGGATATCCCCCGAAGTTCTGAACCGTGTGTACTAGTATGACGCTAGCTG 840
QY 3275 CTTTGAATGTACAGCTTCCATGCGAACAACACCGCTGTGCTATTTCACTGACGCGA 3334
Db 841 CTTTGAATGTACAGCTTCCATGCGAACAACACCGCTGTGCTATTTCACTGACGCGA 900
QY 3335 CGAAACCTGACGCTGCTGGAAGAGAGCTGCAAAATTCAAATTACGACACGCTGTGAG 3394
Db 901 CGAAACCTGACGCTGCTGGAAGAGAGCTGCAAAATTCAAATTACGACACGCTGTGAG 960
QY 3395 AACGCTTGAATGTGCTTCAACGCGCGCCCAACGACATTTACGAGATTTCACTG 3454
Db 961 AACGCTTGAATGTGCTTCAACGCGCGCCCAACGACATTTACGAGATTTCACTG 1020
QY 3455 ACCAGTCCCTACCTGGGCTTCTGACCGTATTGACAGACTCAACGCGGTTTGACGCA 3514
Db 1021 ACCAGTCCCTACCTGGGCTTCTGACCGGATTGACAGACTCAACGCGGTTTGACGCA 1080
QY 3515 ATAAAAATTGAGAACTGTGGAGCGAATCTGATGTGATGATGATGATGATGATGATG 3574
Db 1081 ATAAAAATTGAGAACTGTGGAGCGAATCTGATGTGATGATGATGATGATGATGATG 1140
QY 3575 GCACAAATTCGGCTACATCAGGACAGGACTGCGGATGTCACCAATTCCTGTTACATGCT 3634
Db 1141 GCACAAATTCGGCTACATCAGGACAGGACTGCGGATGTCACCAATTCCTGTTACATGCT 1200
QY 3635 TTGACACGACGATGATGATGAGGAGACAGTATGAGGAAAAATTGATGATGATGATG 3694
Db 1201 TTGACACGACGATGATGATGAGGAGACAGTATGAGGAAAAATTGATGATGATGATG 1260
QY 3695 GGAAGCTGCGCTGCTTGGCCACAAAGGCTACTTCTGTTAGCTCAATGTCCTCAAGT 3754
Db 1261 GGAAGCTGCGCTGCTTGGCCACAAAGGCTACTTCTGTTAGCTCAATGTCCTCAAGT 1320
QY 3755 GACAGTGTACCGCTGATGATGACGAGCGGAGAGCTTGAATTTCACTGACGCTGAGAA 3814
Db 1321 GACAGTGTACCGCTGATGATGACGAGCGGAGAGCTTGAATTTCACTGACGCTGAGAA 1380

QY 3815 AAGATCAGAGGAAGTTTGTGCGTAGAGAGATCTTGTTCACCCGCTCATGAAAG 3874
Db 1381 AAGATCAGAGGAAGTTTGTGCGTAGAGAGATCTTGTTCACCCGCTCATGAAAG 1440
QY 3875 CTGGTAAAGTCCACGTTTACATCATCTTGAAGAGACGCTGCGGGTATCATTAACATG 3934
Db 1441 AGGGTAAAGTCCACGTTTACATCATCTTGAAGAGACGCTGCGGGTATCATTAACATG 1500
QY 3935 CACAGGCGGACCCACACGCTATTAAGCTTATGCTGAGAGAGAGTGTGAGGGAAGTGTAC 3994
Db 1501 CACAGGCGGACCCACACGCTATTAAGCTTATGCTGAGAGAGAGTGTGAGGGAAGTGTAC 1560
QY 3995 ATTAACCACTTCTGCGAAGACGTCACCTACGAATGTAAGTGTGCGACTACAGCA 4054
Db 1561 ATTAACCACTTCTGCGAAGACGTCACCTACGAATGTAAGTGTGCGACTACAGCA 1620
QY 4055 GGTATCGTGAACGCGGACGAGATGAGACGCTGCACTAAAGCAAAACAGTGCATTGCC 4114
Db 1621 GGTATCGTGAACGCGGACGAGATGAGACGCTGCACTAAAGCAAAACAGTGCATTGCC 1680
QY 4115 TACAAGCGGACCAAGGAAATGGGCTTCAACTGCGCGGATCTTATTAAGGACACAGAC 4174
Db 1681 TACAAGCGGACCAAGGAAATGGGCTTCAACTGCGCGGATCTTATTAAGGACACAGAC 1740
QY 4175 CACTCAGTGCAGAGGTTAAATTCACATTCATTCCGCTTGAACACGACAGTCTGCCGCT 4234
Db 1741 CACTCAGTGCAGAGGTTAAATTCACATTCATTCCGCTTGAACACGACAGTCTGCCGCT 1800
QY 4235 CCGTTAGCTCAACGCTTACAGTCAAGAGTGTTCAAAGGATCACCTTCACCTGACT 4294
Db 1801 CCGTTAGCTCAACGCTTACAGTCAAGAGTGTTCAAAGGATCACCTTCACCTGACT 1860
QY 4295 GCAATGCGACCAATGCTGTCGACAAAGGAAATTTGGGCTGCGAGGAGAGGCAACGCA 4354
Db 1861 GCAATGCGACCAATGCTGTCGACAAAGGAAATTTGGGCTGCGAGGAGAGGCAACGCA 1920
QY 4355 GAATGATTTACAGGCTTCACTCAGAGATTTTCTGCGGGCGAGAAAGGCTGAGTAC 4414
Db 1921 GAATGATTTACAGGCTTCACTCAGAGATTTTCTGCGGGCGAGAAAGGCTGAGTAC 1980
QY 4415 GTATGGGTTAACATGAAACGATCAGAGTGTGGCCGACGAGTGTGCGACGACGCA 4474
Db 1981 GTATGGGTTAACATGAAACGATCAGAGTGTGGCCGACGAGTGTGCGACGACGCA 2040
QY 4475 CATGATGCGCGGATGATGATCATCATCATATATATATGCGGATCCAGTCTTACATGCT 4534
Db 2041 CATGATGCGCGGATGATGATCATCATCATATATATATGCGGATCCAGTCTTACATGCT 2100
QY 4535 ATTGCTGTGTGTGTGCTCTGCTATCCTGTGAGGACATGATCATCAGAGCTTGC 4594
Db 2101 ATTGCTGTGTGTGTGCTCTGCTATCCTGTGAGGACATGATCATCAGAGCTTGC 2160
QY 4595 ATGCGCAAGCAGAGAGACTGCGCTGACGCGCATACGCGCTTGCAACGAGCAACGCTA 4654
Db 2161 ATGCGCAAGCAGAGAGACTGCGCTGACGCGCATACGCGCTTGCAACGAGCAACGCTA 2220
QY 4655 CCCACAGATTAAGCGGTTTTGTGCTGATTCGCGCAACCAATGACGAAACATTTGAGAA 4714
Db 2221 CCCACAGATTAAGCGGTTTTGTGCTGATTCGCGCAACCAATGACGAAACATTTGAGAA 2280
QY 4715 ACTTTGAACCATCTGTGTTTTAAACCAACCGTTTTCTGTGGGACAGTGTGATTCCT 4774
Db 2281 ACTTTGAACCATCTGTGTTTTAAACCAACCGTTTTCTGTGGGACAGTGTGATTCCT 2340
QY 4775 CTGCGACGCTTGTATTTCTGCTGCTGCTTTTCACTGCTCATGCTTTTATTTGTT 4834
Db 2341 CTGCGACGCTTGTATTTCTGCTGCTGCTTTTCACTGCTCATGCTTTTATTTGTT 2400
QY 4835 GCAAGCGCTTCTGCGGAGAGTGAAGCGCTTGCAATGAGACCATGTCGCAAAATTT 4894
Db 2401 GCAAGCGCTTCTGCGGAGAGTGAAGCGCTTGCAATGAGACCATGTCGCAAAATTT 2460
QY 4895 CCGGGAATCCGTTATTAAGGCGTTGTGTCGAACGCGCAGGTTACGCGCACTTAACCTGAG 4954
```

```

Db      2461 CCGGGATCCGATTAAGGCGTGTGGAAGTGCAGGTTACGGCCACTTAATCTGGAG 2520
Qy      4955 ATCAGCGTCTCTATGGAATTAAACCTTCACTAACAGAGGATAGCTGACTGCAA 5014
Db      2521 ATTAAGTCTCTATGGAATTAAACCTTCACTAACAGAGGATAGCTGACTGCAA 2580
Qy      5015 TTCACACAGTCAATTCCTTCAACCAAGTTAAATGCTGGGGTCCCTCGAGTGCAGCA 5074
Db      2581 TTCCACACAGTCAATTCCTTCAACCAAGTTAAATGCTGGGGTCCCTCGAGTGCAGCA 2640
Qy      5075 TCCTCAAGGCGGATTACACATGCGCGGTTTGGCGGTGTAACTCTTTCATGTGCGGA 5134
Db      2641 TCCTCAAGGCGGATTACACATGCGCGGTTTGGCGGTGTAACTCTTTCATGTGCGGA 2700
Qy      5135 GGGGCAATATGCTTCTGTGACATGAGAAACAACAATGATGAGGGGTAGTGAATTC 5194
Db      2701 GGGGCAATATGCTTCTGTGACATGAGAAACAACAATGATGAGGGGTAGTGAATTC 2760
Qy      5195 GCTTCAAGTCAATTCCTTCAACCAAGTTAAATGCTGGGGTCCCTCGAGTGCAGCA 5254
Db      2761 GCTTCAAGTCAATTCCTTCAACCAAGTTAAATGCTGGGGTCCCTCGAGTGCAGCA 2820
Qy      5255 GTCGGCTGCTATATGATAGGCAACACACCGCGACTGTGATAGCTTGTCAATGCG 5314
Db      2821 GTGGCTGCTGCTATATGATAGGCAACACACCGCGACTGTGATAGCTTGTCAATGCG 2880
Qy      5315 GTACCGGCAAGTCTTCAACCGGACTTGAAGTCTAGCAGGGCGGATACGCGCTTTT 5374
Db      2881 GTACCGGCAAGTCTTCAACCGGACTTGAAGTCTAGCAGGGCGGATACGCGCTTTT 2940
Qy      5375 TCACCTTTTGCATTAAGTGTGATCAGAAAGGGGCTTGTTCACACTAGCACTTCCCT 5434
Db      2941 TCACCTTTTGCATTAAGTGTGATCAGAAAGGGGCTTGTTCACACTAGCACTTCCCT 3000
Qy      5435 GAGTATGAGCTATGAAACAGAGCGTTCGCGCATATTCAGCATCTCGCTTATGCT 5494
Db      3001 GAGTATGAGCTATGAAACAGAGCGTTCGCGCATATTCAGCATCTCGCTTATGCT 3060
Qy      5495 ACAGACATATGATGCGGCACTGACATACGCTGTGAAGCTTGTTCACAAATCATC 5554
Db      3061 ACAGACATATGATGCGGCACTGACATACGCTGTGAAGCTTGTTCACAAATCATC 3120
Qy      5555 GTCCCTTACACCCAGGATGATGAGGATGAAATGTGAAGAACAACTCGAGCGACCC 5614
Db      3121 GTCCCTTACACCCAGGATGATGAGGATGAAATGTGAAGAACAACTCGAGCGACCC 3180
Qy      5615 CTGCAAGAAACAGCACATTTGATGTAATAATTGAAGTGAAGCTCTGCGAGCGCTTAC 5674
Db      3181 CTGCAAGAAACAGCACATTTGATGTAATAATTGAAGTGAAGCTCTGCGAGCGCTTAC 3240
Qy      5675 TGTGCTTACGGGCAATCCCTATCTCAATTAACATCCCTGATGAGCTTGTGAATCA 5734
Db      3241 TGTGCTTACGGGCAATCCCTATCTCAATTAACATCCCTGATGAGCTTGTGAATCA 3300
Qy      5735 TCAGATCAACCAATTTTGAAGTTAGTGCACAGTAGCAGCTGATTTATTTGCA 5794
Db      3301 TCAGATCAACCAATTTTGAAGTTAGTGCACAGTAGCAGCTGATTTATTTGCA 3357
Qy      5795 GACTTGTGTGTTCTTAACATTAAGTCAAGCTGACAGGAGGAGCAATGTCTCAAT 5854
Db      3358 GACTTGTGTGTTCTTAACATTAAGTCAAGCTGACAGGAGGAGCAATGTCTCAAT 3417
Qy      5855 CACTTCCCACTCAGACAGCTGTTTGAAGAGAGGACACATGTGATGCGGTAGGC 5914
Db      3418 CACTTCCCACTCAGACAGCTGTTTGAAGAGAGGACACATGTGATGCGGTAGGC 3477
Qy      5915 AGCATTAACACTAATTTTAGACATCGAGCCCAAGCAATTTTATGTTTGTGATGC 5974
Db      3478 AGCATTAACACTAATTTTAGACATCGAGCCCAAGCAATTTTATGTTTGTGATGC 3537
Qy      5975 GGCAGAAAGTCACTGCAATGCTGAATGTAAACACCGCGGACCATTAATTGAGAA 6034

```

```

Db      3538 GGCAGAGAACCACTGCAATGCTGAATGAACCAACCGCGCACACATTAATTGAGAA 3597
Qy      6035 CCACATAAGTGCACCAAGAAATTCAGGGCGGAGTTTCAAAACATTTGGAACTGGCTG 6094
Db      3598 CCACATAAGTGCACCAAGAAATTCAGGGCGGAGTTTCAAAACATTTGGAACTGGCTG 3657
Qy      6095 CTTCGACTGTTTGGGAGCATATCCCTCATTTGTTGATGACTTAAATGTTGCTGC 6154
Db      3658 CTTCGACTGTTTGGGAGCATATCCCTCATTTGTTGATGACTTAAATGTTGCTGC 3717
Qy      6155 AGCTTATGCTTATTAACACAGTAGATGACTGAGCGCGGACACTGACATAGCGGTAAA 6214
Db      3718 AGCTTATGCTTATTAACACAGTAGATGACTGAGCGCGGACACTGACATAGCGGTAAA 3777
Qy      6215 CTGATGACTTCCAGGAAACCGGTGTCATTAATGCCAGCGCGGTGACATTAAT 6274
Db      3778 CTGATGACTTCCAGGAAACCGGTGTCATTAATGCCAGCGCGGTGACATTAAT 3837
Qy      6275 CGATGATTTCCAGGAAACAGCATGATATGCTGTGACAGTGCATTAATGATAT 6334
Db      3838 CGATGATTTCCAGGAAACAGCATGATATGCTGTGACAGTGCATTAATGATAT 3897
Qy      6335 CACACTACATATTAACACATATATCATTTTATGAGACTCACTATGCGTCTTAATAT 6394
Db      3898 CACACTACATATTAACACATATATCATTTTATGAGACTCACTATGCGTCTTAATAT 3956
Qy      6395 ACACATACATATTTTAACTTAAACCATATACACTTATTAATCTTTTAAATTTT 6454
Db      3957 ACACATACATATTTTAACTTAAACCATATACACTTATTAATCTTTTAAATTTT 4016
Qy      6455 TCTTTGTTTATTTTATTTTAAATTTT 6485
Db      4017 TCTTTGTTTATTTTATTTTAAATTTT 4047

RESULT 4
AF339476
LOCUS
DEFINITION
Highlands J virus strain B-230 nonstructural protein 4 gene, complete cds.
ACCESSION
AF339476
VERSION
AF339476.1 GI:28193935
KEYWORDS
Highlands J virus
ORGANISM
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus; WEIV complex.
REFERENCE
1 (bases 1 to 4692)
Kinney,R.M. and Pfeiffer,M.
Nucleotide sequence analyses of the 26S mRNAs of viruses of the genus Alphavirus
JOURNAL
Unpublished
2 (bases 1 to 4692)
Kinney,R.M. and Pfeiffer,M.
Direct Submision
AUTHORS
Submitted (23-JUN-2001) Arbovirus Diseases Branch, Centers for Disease Control and Prevention, P.O. Box 2087, Fort Collins, CO 80522, USA
FEATURES
source
1..4692
location/Qualifiers
/organism="Highlands J virus"
/vision
/mol_type="genomic RNA"
/strain="B-230"
/db_xref="taxon:11024"
/country="USA"
/note="isolated in 1960; virus passage P5/SW3/BHK21-1"
<1..564
/codon_start=1
/product="nonstructural protein 4"
/protein_id="A03332.1"
/db_xref="GI:28193936"

```

translation="FKFGAMKMSGNFLTFVNTLVNIMIASRVLEERLTTSACAFIG
DNIIVGVSDNIMADRCATWLNMEKIIDAVIGKAYPCGFLVIOIGACRVA
DPLKRLFKLGLPLDDODVDRRALDESLRMRIITSELVAVASREVLVLSL
ITIALSLADSVKFKIRGNPILYX"
565..4692
/product="structural polypeptide precursor"
/note="268"
565..605
606..4316
/codon_start=1
/product="structural polypeptide precursor"
/protein_id="AA033323.1"
/db_xref="GI:28193937"
translation="MFYPOLTFPMYQPNMAVRDPPRRMRPFRPLAOTEDL
RRIANLTKORAPNPAGPPAKKKKTPKXAPPKKKKQOVKKKKPKRGROR
CMKLESDTPILLNGVNYACVGVGLMRPHVEGKIDNEQLAIIYLKASWIDLE
YGDVPONKSDILOITSEKPPFTNMHGAQVYENGRTFVRGCKDSRPLDNK
GRVAIVLGKSNESRTALSVTNQKVTIKDPEGEQMSLTAMCVLANVFPD
KRPVCSLAPSRITLVEENDPQYDITLNLVKPSRQKRSITMDPTLSPYLGH
CPYCLHATPFSPIKIEVWDESDGTIRIQVSAQGNQAGTADPTFRVMSYEDH
DIKESMDKIASISGPGSRLGHKGFIILARCPDSDVTSVTSQASATSCVEKIR
KPFYGREBYLLPVMHGKIKCHVDHLEKTAGYITMRPAPHATATVSESSGVYI
RPSGKNTYRECKGSDISTGTNTRTKMPEGCKKQCIAYRHDTKWVNSPDLIRHS
DHAVQKLIIFKLTALACVPLAHTPVEKMFKQVTLALASHPTLLTRKLGPRAR
PTSEWIVGTVRNFSVREGELEYTMGNMDPVKWSQESAPDPMHPIIIVHYHRH
PLTYIAVLGGLVLTIVIGIASAAICISKARDCLTPYALAPNAVPTLVAICIRPT
HAETLGEISGLHMLNNOPLMAOLCLPALAILLRFPSCLPELVAVGLGADAY
EHATTVPVPGVPYKALYERAGYALNLEVYVSESLPSTNKEVTCRHTIIPSPQ
VKCCGSLCCQSRKADYCRVFGVYPPWKGAGCFQSENTQJSEAVIEAPDCTAD
HAVALKHTAALKVGLQIVYGNSTRIDTFVNGVPGSSRLKALVISAFTPEFDH
KVIRKGVNYNDPEPEYGMKRGVFDIQASLSDSTDIARBDIRLSPVKSIIHVPY
TOAAGYEMKNNNSGRPLQDTPARFCCKLEVPDLRAVDCAVGHIPISIDIPAAFRYS
EAPYIEMSCVTYACTIVADFGSLTLOYKDKRGNCHVSHSSNAVLKEATYVHS
GSVILHFTSSPQANFYISLCGKTTCAECKPESDHIIGEPHKVNOEFOAAVSKTSH
NWLFAMLGASSLIVGLVLVACSSMIINTSR"
mat_peptide
/product="capsid protein"
1383..1442
/product="E3 protein"
mat_peptide
1443..2831
/product="E2 envelope glycoprotein"
2832..2996
/product="E6 protein"
2997..4313
mat_peptide
/product="E1 envelope glycoprotein"
4317..4692
3'UTR
BASE COUNT 1324 a 1183 c 1105 g 1080 t
ORIGIN

Query Match 41.7% Score 2705.6; DB 14; Length 4692;
Best Local Similarity 75.6%; Pred. No. 0; Mismatches 1084; Indels 4; Gaps 2;
Matches 3380; Conservative 0;

QY 1866 TTTAAATTGGTGCATGATGAATCCGATATGTTCTTAAACGCTGTTTCAACACACTA 1925
DB 1 TTTAAGTTCGAGGCATGATGAATCAGGAATGTTTTGACCTGTTTGTGTAATCTCTA 60
QY 1926 GTCAATATCATGATTGCTAGCAGAGTACTAGCTGAACGTTTAAACGCTCAGCGCCG 1985
DB 61 GTTAAACATCATGATGACAGTCGGGTACTTAGAGAAAGTTGACCACTTCAAGCGTGTCA 120
QY 1986 GCTCTATCGGAGCATATACATAGTGTGTCGTCTCCGACACCTTATGCGGAG 2045
DB 121 GCTTTATCGAGAGCAATATATGATACAGGTGATGATCCGATATTTGATGCTGAT 180
QY 2046 AGATGCGGCACTTGCTGAACATGAGAGTAAATATATGATGACGTTATTTGATCAAA 2105
DB 181 AGGTGTGCAGGTGCTAAATATGAGATAAAAATCATTTGATGCCGTATGCGCTAATA 240
QY 2106 GCAACCTTCTTGTGGGGGATTTATCTGTGTGACCGATGATACGAGCAACGCTGCA 2165
DB 241 GCACCATCTTTTGGGAGGGTTCATCTGTGTGATCAAAATTAACAGTACAGCGTGTAG 300

QY 2166 GTCCGAGACCCCTTAAAAAGGCTTTTAAAGCTGGAAAAACATTGCCAGTGATATACC 2225
DB 301 GTAGAGAGCCCCCTCAACCGTTTGTGTAACGTGGTAACACTCCCTTTAGACATGAT 360
QY 2226 CAAGCTGCGACCCGCGGCGACCTGATATGAAAGCAATGCGATGGAACGAATTGCA 2285
DB 361 CAGATGTAGACCCCGTCGAGCGTTCGACATGATGCTTAAAGGTGAATATGATAGT 420
QY 2286 ATTACGAGCGATTAGTGAAGCCGTGAAATCCAGATTCOGAGTATCTGCGAGCCTG 2345
DB 421 ATAACTCTGTGACTTGTGTAAGCGGTGAAATCCGATVYAGGTGTCTGTCTAGCTTA 480
QY 2346 ATCATCAGTCTGTGTCCAGCTTACGCCGAAGCGTTAAGAACTTCAAGAGCATAGAGG 2405
DB 481 ATTATTTATGGCTGTCAACACTTGCAGATAGCGTGAAAGAACTTTAACCGATAGAGA 540
QY 2406 AGCCCAATCACCTCTACGCGTGAACCTAATAGTGAACGTATGACACGCACTAACCA 2465
DB 541 AACCCATTACTCTTACGCGCTGACCTTAATAGTGAACGTATGATACGCACTTA---A 597
QY 2466 CCGGCAAGATGTTTCCATACCCCTCAGCTGAACCTTCCACAGTTTACCTTAACAAATCCGA 2525
DB 598 CCGGCAAGATGTTTCTTATCCGCAATGACCTTTTCCACTATGTATGAACCCAAATCCGA 657
QY 2526 TGGCTTACCGAGATCCAAACCCCTCCTAGGCGCGCGTGGAGCCGTTTGGCGCCCGCTGG 2585
DB 658 TGGCATACAGAGACCCGAACCTCTAGCGCCAGTGAAGCCATTAGCGCCCACTTG 717
QY 2586 CTGCTCAATGGAAGATCTTAGAGAGTGCATATGTCATCTTCAACACGATCAC 2645
DB 718 CAGCACAAAATAGAGACTTGGAGAGTGCATCGCAAAATTTGACCGTTAAACAGAGAGAC 777
QY 2646 CTAATCCGCGCCGAGTCCACCCGCAAGAAAGAAAGAGTCTCTTAAGCCAAACCTA 2705
DB 778 CCAATTCACCTGCGCGTCTCTCACTTAAGAAAGAAAGAACTCAACCGAAACCGAAGGCA 837
QY 2706 CTCAGCCTTAAAAAGAGAGCAGCAAGCCAGAGGCAAGCAAGCCTTAACCCAGGGA 2765
DB 838 CGCCACCTTAAGAAAGAGAGCAGAGCTTAAGAAACAAACCTTAACCCAGGGA 897
QY 2766 AACGACAACTGATGTATGTAAGATTGAGTGGAGCAAGACATTTCCGATCATGTGAA 2825
DB 898 AAGGCAACGATGTATGTAATTTGAAATCGAACAAACCTTCCGATCTGTGTAACG 957
QY 2826 GCCAAGTGAATGATATGCTGCTGCTGTCGAGAGAGGCTGATGAACCACTCCACGTTG 2885
DB 958 GGCAGGTAAATGCTACGCAATGCGCTGTTCGCGAGGCTATGAAACATTGCAATGTGCG 1017
QY 2886 AAGGAAAAATGATATGAGCAATTAAGCGCGCGTGAATTAAGAAAGCTAGCATGTAGC 2945
DB 1018 AAGGTAATATGATTAAGCAACAGTTGGACGATCAATCAATGAAGAAAGCAGTATGACG 1077
QY 2946 ACTTGAATGAGGAGCGTTTCCAGAAATGAAATGACAGCGTGCAGTACACAGCG 3005
DB 1078 ATTCGAGTACGAGTACGACGTAACCCAGAAATGAATGACAGACCTGCAATACACAGCG 1137
QY 3006 ACAAACCAACCGGCTTCTTACCACTGSCACCAACCGCGCATGATGTAAGATGGAAT 3065
DB 1138 AAAAAACCACTTGATTTCTATTAATGCGCATCTATGCGCATGATGTAAGATGGAAT 1197
QY 3066 TTACCGTACGAGAGAGTGGCGGGAAGAGCGACAGCGGAAGACCGATCTCTGACACA 3125
DB 1198 TCACCGTTCCACGCGGAGTCCGTGTAAGAAAGAGACAGTGGAGAACCTTCTGACACA 1257
QY 3126 GAGGCAAGTGTGCTATTTCTTGAAGGTGCAATGAGGCAACGCTTACGCGCTTT 3185
DB 1258 AAGCGCGGATGTGCTATTTCTTGGAGGGGGAAGAAAGTTCAAGGACGCGCTGT 1317
QY 3186 CAGTGTCTACTTGAACCAAGAAAGGGTACCATTTAGGATATCCCGGAAGTCTTGAAC 3245
DB 1318 CGTGTGTTACATGGAACCAAAAGGGTGAACATCAAGGAATCCCTGAAAGCGCAAC 1377
QY 3246 CGTGTCTACTGTTAAGACGCTATGCGTGTGTTGAAATGTACAGTTCATGCCAGCAAC 3305

Db 1378 AATGTCATTGATCACTGCGAATGTCGTCCTTGCAAAATGTCATCTTCCCGTGTGCAAAAC 1437
Qy 3306 CACCGGTGTGCTATTCCTGACGCCGAAACGAACTGCGACTGCTCTGGAGAAAGCTCG 3365
Db 1438 CTCTGTCTGTCTATTCCTGTAGCACAAGAACGCACTTGATGTCTGGAGAAACCTGG 1497
Qy 3366 ACAATCCAAATTAGACACGCTGCTGAGAAAGCTTGAAATGTCCAAATCCGCGGCCA 3425
Db 1498 ACAATCCAGGTTAGACACCTGCTAGAAATGTACTGAAGTGCCTCCACGCTGACAGA 1557
Qy 3426 AACGACATTACCGATGACTTTCACACTGACAGTCCCTACCTGCGGTTCTGCCGTAAT 3485
Db 1558 AAGGACATTAGGAGAGATTTTCACTTGACAAAGCCGTAATTAGGCAACGTCGTAAT 1617
Qy 3486 GCACACTCAACCCCGTGTTCAGCCCAATAAAAATTGAGACGTGTGGACAAATCTG 3545
Db 1618 GCCTCCACGCTACCCCTGCTTGTAGTCCATTTAAATTTGAAACGTTTGGGACAGTCAG 1677
Qy 3546 ATGATGATGATTAAGATTCAGGCTCGGCAAAATTCGCTACATTCAGGACAGGCACTG 3605
Db 1678 ACAGCGTACATAGAAATACAGTTTCGCGCAATTAAGTTACAAACGACAGATACGG 1737
Qy 3606 CGGATGTACCAAAATTCGTTACATGTCCTTTCGACACAGCAATGACATCAAGGAGACA 3665
Db 1738 CAGACCAACTTAAGTTCAGGTACATGTCCTACGACAAAGACATGACATCAAGAAAGTA 1797
Qy 3666 GTATGAGAAATATGCTATCAGCAATCTGGAACCTCGCGCTGTCTGGCCACAAGGCT 3725
Db 1798 GTATGAGCAAGATGCTATCATGTAATCTGGGCGGTGACGCGCTGGGCAATTAAGCT 1857
Qy 3726 ACTTCCTGTTAGCTCAATGTCCTCAGGTGACAGTGTACCGGTACAGTATCAAGCGGAG 3785
Db 1858 ATTTCTGTGCGCAGATGCCACAGAGACAGCGTAAACGTGACATCAACGATGAG 1917
Qy 3786 CATCTGAAATTCATGACACGCTGAGAAAAAATCAGAGAGAAATTTCTCGGTAGAGAG 3845
Db 1918 CATAGGAGACATCTGTGACGCTGAGAAAGAAATTAAGAGGAAATTTGTGGGAGAGAG 1977
Qy 3846 AGTACTGTGTTCCACCCGTCATGAAAGCTGTAAAGTCCACGTTTACATCATCTTGA 3905
Db 1978 AGTACTGCTCTCCACAGTGACGCGCAAGCTAACTCAATGATATATACCACTTAA 2037
Qy 3906 AGGAGAGCTGCGGGGTACATAACATGACAGGCGGCAAGCGGTATTAAGTCT 3965
Db 2038 AGGAGACAAAGAGATATTAATATGACACAGGCGCAATGCGTAAACCACT 2097
Qy 3966 ATCTGAGAGAGCGCTCAGCGAAAGTGTACATTTAAACCACTTCTGCGAAAGACGTCACT 4025
Db 2098 ACGTGAAGAAATCATTCGGGCAAGGTATCATCCGTCCACTTCGGAAAGAAAGTGAAGT 2157
Qy 4026 ACGAATGTAAGTGTGGGACCTACAGCACAGGATCTGTGAGACCGGCAACGAATGAACG 4085
Db 2158 ACGAATCAATGTGTGAGATTACAGCACCGGACAGTTTAACTAAGATATTAAGATGCGAG 2217
Qy 4086 GCTGCTAAAGCAAAACAGTGACTTGTCTACAGAGGCAACAAACGAATGGGTCTCA 4145
Db 2218 GATCAGAGAAAGAAACAGTGACTTGTATTAACAGACCAACCAATGGGTCTTTTA 2277
Qy 4146 ACTGCGCGATCTTATTAAGGACACAGACCACTGACAGTGAATTTGCAATTCAT 4205
Db 2278 ATTACACAGACCTTAATTCGTACTCAAGCAATGACAGTGAAGAAACTGCAATTCCT 2337
Qy 4206 TCGCGTTGACACGCAAGTCTGCGCGTTCGCTTAACTGACACGCGTACAGTCAAGT 4265
Db 2338 TTAATACTTAAGTCTACGCGCATGCCCAGTGCACCTCGCTACACACCAACGCTTGAAGT 2397
Qy 4266 GCTTCAAGAGATACCTCCACTGCACTGCAATGCGACCAATGCTGTGACAGAGAA 4325
Db 2398 GGTTCAAAGGCGTACGCTGACCTGACCTGACTGATGATCCACACTGTTGACTAGCGGTA 2457
Qy 4326 AATGGGCTGCGAGACAGCAACAGCAGATGATTAAGGCTCTCATCCAGAAAT 4385

Db 2458 AGCTAGACCGGAGCTGAACCCACGCTGAATGATCTGTTGGAGCGCTGCCAGAACT 2517
Qy 4386 TTTCTGTGGGCGAGAAAGGCTGAGTACTATGGGTAACATTAACCATGTCAGAGTCT 4445
Db 2518 TCTCCGTAAGGAAAGGAGGCTCGAATATATCTTGGGAAATTCACGACCCGTCGAGTGT 2577
Qy 4446 GGGCCGAGAGTGGGACACAGGCGACCAATGATGAGCGGCTATGATCATCCACT 4505
Db 2578 GGTCTCAAGAGTGTGACCTGCGGAGCCCTACGCTTGGCGCACAGAAATATGTAAT 2637
Qy 4506 ATTATCATGCGCATTCAGTCTACATCTGTATGCTGTGTGTGCTCTTGTCTATTC 4565
Db 2638 ACTACCATCGGCAATCCGCTATTAACCAATGCGGTGTGTGCGGACTGTCTAATAACAG 2697
Qy 4566 TGTAGGCACTGATCATTCAGAGCTTGTGATCCGCAAGCAAGAAAGAGTGCCTGACG 4625
Db 2698 TGAATGCGATCGCTTTCGACGCGGCTGCAATATGAAAGCGGAGAGATGCTGACAC 2757
Qy 4626 CATACGCTTGGACGAGACGCAACGTAACCAAGCATTAAGCGGTTTGTGCTGACATTC 4685
Db 2758 CTAACGCTTGGACCAACGCTGCACTTCCAAGTTGCGGAGTCTTATGCTGACATA 2817
Qy 4686 GGCACCAACCGCTGAACATTTGAGAACTTTGAACATCTGTGTTTAAACCAAC 4745
Db 2818 GACCTACGACGCGGAGACCTAGGCGAATCACTTGGACACTCTGCTTAAACATTAAC 2877
Qy 4746 CGTTCTCTGGGACAGTTGTGATCTCTGACAGCGCTGTGTTATCTGTCCGCTGCT 4805
Db 2878 CTCTTTATAGGCGAGCTGTGCTTCACTTGTCCGCACTTATATCTTTCAGGTTCT 2937
Qy 4806 TTTCACTGCTCATGCTTTTATATGTTGTCAGAGCGCTCTGCGGAAAGTATGACGCT 4865
Db 2938 TCTCATGCTCTCTCTCTTTTATGTTGTGACGCGCTCTGCGGAAAGGACGAGCTCT 2997
Qy 4866 TCGAACAATGCGACCATGTCGCAAAATGTTCCGGGATCCCGTATTAAGCGCTGTGAC 4925
Db 2998 ACGAACAATGCGACCATGTCGCAAAATGTTCCGGGATCCCGTATTAAGCGCTGTGAAA 3057
Qy 4926 GCGGAGTTACGGGCACTTAACCTGAGATCAAGGCTGTCTCATGCGAATTAACCTT 4985
Db 3058 GGGCGGGATGACCCCTTAACCTGAGATCAAGCTGTGTGCTGAACTGATACCT 3117
Qy 4986 CAACTAACAAGAGTATGTAAGTCTGCAAAATTTCCACAGTCAATTCCTTCCACCAAGTTA 5045
Db 3118 CAACAATTAAGATATGTAAGTCTGCAAAATTTCCATACATTCATCACCCAGGTGA 3177
Qy 5046 AATGCTCGGGTCTCTGAGTGAAGGCAATCTCAAAAGCGGATTAACATGCCGCTT 5105
Db 3178 AGTGTGCGGTCCTTGAATGTCAAGCTCAAGCAAGGACGACTATACCTGTGCGCTGT 3237
Qy 5106 TTTGCGGTGTGTACCTTCAATGTGGGAGGCGCAAAATGCTTCTGGAACAGTGAACA 5165
Db 3238 TCGAGGTGTCTACCAATTAATGTGGGAGGCGCAAGTCTTCTGGAATGAGGAAACA 3297
Qy 5166 CACAACGATGAGGCGCTAGCTGAGTTCGCTCAGACTGCACTATAGATCAAGCACTG 5225
Db 3298 CGCACTGAGAGAGGCTATGTTGAATTCGCTCAAGCTGACACAGCAATCAAGCACTG 3357
Qy 5226 CACTAAAGTTTCAACAGCTGCTGTGAAATCGGCTGTGTATATATACGCAACCA 5285
Db 3358 CACTGAAGTCTCACACAGCTGCTGTGAAGGCTGCGCTCAAAATGATATATGAGGAAACAT 3417
Qy 5286 CCGGCACTGATACCTTGTGAAATGCGCTCAAGCGAGTTTCTCAACGCGGACCTGAGG 5345
Db 3418 CAACACGCTGGAATCAATTTGTCAATGAGGATTAACACAGGCTCATCAACGCACTGAAAG 3477
Qy 5346 TCAATGAGGCGGATTAACAGCGCTTTTACCTTTGACCAATTAAGTGTCTACAGAA 5405
Db 3478 TGAATGAGGCGCTATATCAGACGCTTACACCTTTGACACCAAAAGTGTCTATTCGAA 3537
Qy 5406 AGGGGCTGTTTTAACTACGACTTCCCTGATATGAGAGCTATGAACCAAGAGCGTTG 5465
Db 3538 AAGGTAAAGTGTACCACTGACTTCCCAAGATACGGTGTATGAAGCCAGGCGGTGTTG 3597

QY	5466	GCGATATTCAGACATCCCTCGTTGATGCTACAGACATAGTAGCCGCACTGACATACGCG	5525
Db	3598	GGGATATCCAG3CATCTTCCTCGGTGATTCACAGACATGCTGTTCAGAACTGACATCCGAC	3657
QY	5526	TGCTGAAGCCCTTCCTGTCAGAAACATCCAGCTCCCTTACACCCAGACAGTATACAGGTTATG	5585
Db	3658	TACTCTAAACCATCGGTAAAGAGCATTTCACTGACCGTACACACAGGCCGCGTCAAGTTACG	3717
QY	5586	AAATGTGMAAACAACACTCAGGACGACCCCTGACAGAAACAGCACATTTGGATGTATAA	5645
Db	3718	AAATGTGMAAATATATTCGCCGAGAACCTCTGACAGATACAGCCCTTTTGCGTCGCAAAA	3777
QY	5646	TTGAAAGTGGAGCCCTCTCGAGCCGCTTAACTGTGCTTACGGGCACTCCCTTATTCGATGT	5705
Db	3778	TCGAGGTAGATCCGTTAAGGCGGCTGACTGTGCTTAAAGGCGCATATCCCGCTCCACATCG	3837
QY	5706	ACATCCCTGATGACGCTTTGTGAGATCATCGAATCCAAACAATTTAGAGTTAGCT	5765
Db	3838	ACATATCCGAGACGACGATTCGTTGAGAACCTCTGAAGACCCACCGTGTCTGAATGAGTT	3897
QY	5766	GCACAGTAGACAGACTGACATTTATTTCTGAGACTTTGTGTCTCTTCAATTAAGTAC	5825
Db	3898	GCAAGGTAGCGGCTCGCATTTTACTCCGCAATTTTGTGTGATCTTTAACCTTACAATPAC	3957
QY	5826	AACTGACAGGAGGAGCATTTGCTCAGTTCACTCCACTCCACGACGACGCTGTTTGAAG	5885
Db	3958	AAGCAGACACAGGAAGGCATATGTCTGTCCATTGCGACTCTCTTACGCGGCTGTAAG	4017
QY	5886	AAGGACACACATGTGACTGCCGTAGGACATPACATCATTTTAGCATATCGAGCC	5945
Db	4018	AAGCTACCCACCAGTAGTTACTCTCGGGGAGCGTTACTTTTACATTCACGACGTCAGCTC	4077
QY	5946	CACACAGCAAAATTTATAGTTTCGCTATGCGGCAAGAAAGTCCACCTGCATCTGATGATG	6005
Db	4078	CGCAGGCAAACTTTATTTGTGTCTGCTATGCGGTAAAGAAAACAACATGACGCGGAGTGT	4137
QY	6006	AACCAACGCGCGCAGCACATAATTTGAGAACACATAAAGTGACCCAGAAATTCGAGCGG	6065
Db	4138	AGCCCCCATCGACACATCATATAGGGGAACACATPAAAGTTAACAGAAATTCAGAGAG	4197
QY	6066	CAGTTTCCAAAACATCTTGGAACCTGTGCTGCTTGCACTGTTGGGGGAGCATCATTCCTCA	6125
Db	4198	CAGTTTCCAAAACATCATGGAACCTGCTATTTGCGCATCTGTGGGGGAGCATCATTCCTCA	4257
QY	6126	TTGTTGTAGAGCTATATAGTGTGTGCTGACGCTGATCTGATCTTATPAAACACAGTAGATGAC	6185
Db	4258	TTGTTGTAGAGCTATATAGTGTGTGCTGACGCTGATCTGATATPAAACACAGTAGATGAC	4317
QY	6186	TGAGCGCGGACACTGACATAGCGGTAA-ACCTGATGTACTTCCGAGGAAACGCTGTGCA	6244
Db	4318	CCGGCGTTCCCTCTACCAACGCGGGAAGCACTCATGTATCTTCCGAGTAACTGTGTGCA	4377
QY	6245	TAAATGCCACGCGCGCTTGACACTPAAAACTCGATGTATTTCCGAGGAAGCACAGTGACATA	6304
Db	4378	TAAATGCCACGCGCGCTTGACACCAAAACTGATGTACTTCCGAGGAAGCACAGTGACATA	4437
QY	6305	ATGCTGTGAGGTGCACATTAATGTAT 6332	
Db	4438	ATGCTGTGAGGTGTGCATTTAACCAAT 4465	

RESULT 5	AR042411	AR042411	Sequence 7 from patent US 5811407.	11663 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	AR042411						
DEFINITION	AR042411						
ACCESSION	AR042411						
VERSION	AR042411.1		GI:5962907				
KEYWORDS							
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 11663)						

Query Match	Best Local Similarity	Score	DB	Length
Matches 3752; Conservative	64.6%; Pred. No. 0; Mismatches 1984; Indels 74; Gaps 9;	37.0%; 2397.6; 11663;	6;	
580	ACTGACGGTATGAAACGGGAGCGGTATATTTCTCATCGGAAACAGCCCAAGTCACCTTC	639		
5707	ACTGCTTACACCGGGGTAGTGGGTACATATTTTCAGCGACAGAGCCCTGGGACCTTC	5766		
640	AACGAAATCATGTAACGTCAATGTAACTACAGAACTATTTGGATGGGCGCTCATG	699		
5767	AAAAGAGTCGGTCTCGACAAACGCTTACAGAAACGACCTTGGAGGCAATGTTCTGG	5828		
700	AGAAGTATTACGCCCGCGCTCGATCTCGAAAGAGAAATGTTACAGAAAGAACTCG	759		
5827	AAAGATCTTACGCCCGCTCGATCTCGAAAGAGAAACAGCTCAACTCAGATACC	5886		
760	AATTATGCGCCTCTGAAAGAAATAGAACAGGTATCAATCAAGAAAGTAGAAATATGA	819		
5887	AGATGATGCCACCGAGGCCAAAGACAGGTACAGTCTCGAAAGTAGAAATACCGA	5946		
820	AAGCAATTACAGCGGAGGACTATTTCTGGATTGGGCAATATCTATCATCAGAGTGA	879		
5947	AAGCCATACCACTGACGACGACTGCTTCAGGGCTACGCGTGTATACCTGCGCACAG	6003		
880	ATCCTGTCAGTGTACAGAGTCAATATCTGTACCAATCTACTCGTCAACGGTATTA	939		
6004	ATTCAGCCAGATGTCTATTAAGTACCTACCCGAAACCATGCTATTTCCAGCAGTACAG	6063		
940	ACAGTATTACATCTGCAGAGTCCGCGTTAAACGTCCACTTATGTTATCCAAAGAAAT	999		
6064	CGAATCTCTGACCCCAAAGTTTCTGTAGCTGTTTGTACAACTATCTGATAGAAAT	6123		
1000	ACCCTACAGTACGCAATTATTTGATTAACAAGTAAATCGATGCGTATCTTGACATGCTGG	1059		
6124	ACCAGCGGTAGACTTATTAAGATCAACGACGATGAGATGCTTACTTGGATATGTTAG	6183		
1060	ACGGCGCATGCTGCTGTCTAGATACAGCCACTTTTTCGCGCTTAACTGAGAACTAC	1119		
6184	ACGGGACAGTCTGCTGTCTAGATCTCAACTTTTTCGCGCCCAAGCTTAGAAGTTACC	6243		
1120	CAAGGAAGCATAGCTATTTTGAGGACAGAGATTAAGATCAGCGCTCCCATTCGCTTACGA	1179		
6244	CGAAAGACACGAGTATAGAGCCCAACATCCGAGTCCGCTTCAACGAGATGACGA	6303		
1180	ATACTATTACAAATGATTTGGCTGACACTACTAAAGAAATTTGCAGCTTACCCAAATGC	1239		
6304	ACAGTGTGCAAAAGTGTCTCATTTGCGCGACATAAAGAACTGCAAGTCCACAAATGC	6363		
1240	GAGAAATTACCTGTCTTATAGATTCGGCGCAATTTAAATGTTGATTTTCAAGAAATACGAT	1299		
6364	GTGAAGTCCCAACACTGAGACTCAGCGCAATTTCAACGTTGAATGCTTTCGAAATATGCAAT	6423		
1300	GCAATGATGAGTACTGGGATACCTTTTGGCATTAACCTATTCGCGCTTAACTACAGAAACG	1359		
6424	GCAATGACGATATTGGAGAGATTTTCCCGAAAGCCAAATTTAGGATACATACTGAGTTGC	6483		
1360	TTAGCAATATATGTGACAAAGCTGAAAGGCGCGAAGCAGACAGACTTTTGGCAATATCTC	1419		
6484	TTTCCGATATGTTGCGCAGACTGAAGAGGCTTAAAGGCGCGCAGCACTTTTGGCAAAAGACG	6543		
1420	ATAATCTTAAAAAGCTTGCAGAGATACCAATGATCAATTTGCTATGATCTTAAAGAGAG	1479		
6544	ATAATTTGTGCTCCCTTGCAGAAAGTGTCTTATGATTAATTCGTATGATGAAATGAAAGAG	6603		

QY 1480 ATGCAAGTTACTCCCGGACGAAACATACAGAGGCGGCTTAAGTGCAGTTATTC 1539
DB 6604 ACCTGAAAGTTACCTCGGACGAAACACAGAGAAAGACGAAAGTCAAGTGTATAC 6663
QY 1540 AGGCTGCAGATCCCTTCTTCTACCGCTTACCTTTCGGGATCCATCGGAAATTAGTCCGTA 1599
DB 6664 AAGCCGAGAAACCCCTGGCGGCTTACCTATCGGATCCACCGGAGTTAGTGGCA 6723
QY 1600 GACTGAATGGGCTCTTCTGCAATATCACTACTCTTTCGACATGTACGCGAAGATT 1659
DB 6724 GGGTTACAGCGTTTTCCTACCCAACTTCAACGCTTTCGATGTGCGCGAGACT 6783
QY 1660 TTGATGCAATTTTGTCTGAACTTTCACACGCGGACCCAGTATTGAAACGCAATCG 1719
DB 6784 TTGATGCAATTTGCAAGAACTTCAAGCAAGGTGACCGGATCTGAGAGCGATATCG 6843
QY 1720 CGTCTGTTGATTAAGGCAAGACGAGCTATCGCATTTTCGGCTTGATGATCTTGAAG 1779
DB 6844 CTTCTGTTGCAAAAGGCAAGACGAGCTATGGCTTAACCGGCTGATGATCTTGAAG 6903
QY 1780 ACTTAGGTGACCAACCGCTCTTGAATTTGATAGAGCGGCTTCGCAATATCAT 1839
DB 6904 ACCTGGGTGACCAACCACTACTCGACTTGTGAGTGCAGTGCCTTTGAGAAATATCAT 6963
QY 1840 CTGTGACCTTACCTACGAGACGAGTTTAAATTTGGTGCATGATGAATTCGGTATGT 1899
DB 6964 CACACCACTGTCCACGAGTACCGCTTCAAAATTCGGGCGATGATTAATTCGGAATGT 7023
QY 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGTAGCAGATCTACGTG 1959
DB 7024 TCTCAACGCTTCTTGTCAACACAGTTCTGAATGTCGTTATTCGACAGATATTTGAGG 7083
QY 1960 AAGGTTAATCAAGTCAAGCTGCGGCTCTATTCGCGACGATTAATATGATGATGTG 2019
DB 7084 AAGGCTTAAACGTCMAATGTGCACTTTATTCGCGACGACMAATTAATACAGGAG 7143
QY 2020 TCGTCTCCGACACCTTGAATGCGGAGATGCGCACTTGGCTGAACATGGAAGTAAATA 2079
DB 7144 TAGTATTCGAAAGAAATGAGTGAAGGTGTCGACCTGCTCAATGGAAGTTAAGA 7203
QY 2080 TTAATGATGATTAATTTGATATCAACACCTTACTCTGTGGGGAATTAATCTGTGTG 2139
DB 7204 TCAATGACGAGTATGCGGAGAGACCACTTACTCTGCGGTGATCATCTTGAAG 7263
QY 2140 ACCAGATTAACAGGACAGCTGCAAGTGCAGACCTCTTAAAGGCTTTTAAGCTTG 2199
DB 7264 ATTCTGTTACTCTCAACGCTGTGCGTGGCGGACCTTGAAAGGCTTTTAAGTTGG 7323
QY 2200 GAAACCAATTGCGAGTGAATGATACCGAAGCTGCGACCGCGCGGGCACTGCATGATG 2259
DB 7324 GTAAACCGCTCCAGCGGAGTGAAGCAAGACGAAAGCGGCTCTGTGTATAGT 7383
QY 2260 AAGCAATGCGATGGAACAGATTGGAATTAACGAGTGAAGTGAAGCGCTGATATCA 2319
DB 7384 AAACAAAGGCGTGTGATTAAGTGAATTAACAGACCTTACAGATGCGCGTGAATCTC 7443
QY 2320 GATACGATCATATGCGGAGGCTGATCATCACTCTCTGTCAAGCTTACGCTTAAGCGG 2379
DB 7444 GGTATGAGTGAACACATGACACTGTCTGTGCGATTTGAGAACTTTTGGCCAGAGCA 7503
QY 2380 TTAAAGACTTCAAGAGATTAAGAGGAGCCCAATCACCTCTAGCGGTGACTTAATAGG 2439
DB 7504 AAAGAGCATTTACACCATCAAGGAGGAAATTAAGCATCTCTAGGAGTGTCTTAATAGT 7563
QY 2440 TGAAGTATGA-----CAGCAGCTTACCAACCGGAGAAATGTTTCA 2482
DB 7564 CAGATATGATCAATTTCACTGATTAATACCAACACCAACCATGATTAATAGGATTC 7623
QY 2483 TACCTTACGCTGAATTTTCAACAGTTTACCTTAACAATTCGATGCTTACCGAGATCA 2542
DB 7624 TTAACATGTCGCGCGCGCTTCCAGCGCCCACTGCTATGTGAGGCGCGGAGAA 7683

QY 2543 AACCTCTAGGCGCGCTGAGGCGGCTTTCCGCCCCCGTGGCTCTCAATCGAAGT 2602
DB 7684 GAGGAGAGGCGGCGCGGATGCTGCGGCAATGGCTGGCTTCCAAATCGAACATGTA 7743
QY 2603 CTTAGAGGTGATGATCACTTGAATTAACAA--GATTAATCTTATCCGCGCAG 2660
DB 7744 CCACACCGCTCAGTGGCTTATGATTTGACAGGCACTGACCTCAAAACCCCGCCAC 7803
QY 2661 GTCCACCGCAAGAAAGAAAGATGCTCTTAAGCCAAACCTTACAGCTTAAAGAA 2720
DB 7804 GCGCGCGCGCGCCGCAAGAAAGCAGGCGCAAGGACCAACCGGAGCCGAGAAACCA 7863
QY 2721 AGAAGAGCAAGCAAGAGAGCAAGAAC---GCAAGCTTAAACGAGGAAACGAAACGTA 2777
DB 7864 AAACACAGAGAAAGAAAGAAAGCAACCTGCAAAACCCAAACCCGAAAGAGCAGGTA 7923
QY 2778 TGTGATGAAGTTGAGTGGACACAAATTTCCGATC--ATCTGAACGCCAAGTGA 2834
DB 7924 TGGCACTTAAGTTGAGGCGGACAGACTGTTCAAGCTCAAAATGAGGACGAGATGTA 7983
QY 2835 ATGATATGCTGCTGCTTTCGAGGAAAGCTGATGAACCACTCCAGTTGAAGAAAA 2894
DB 7984 TCGGAGACGACTGCGCATGAGAAAGGTAATGAACCACTCCAGTGAAGAACTA 8043
QY 2895 TTGATATGAGCAATTAGCGGCGGTGAATTAAGAAAGGCTAGCATGATGACTTGAAGT 2954
DB 8044 TTGACCAACCTGTCATTAAGCTCAAAATTCACAAAGTGTGAGCATACGATGAGAT 8103
QY 2955 ACGGCAAGTTCCTCCAAATGAATCAGACAGCTGCAATCACCGAGCAACAAACAC 3014
DB 8104 TCCCAAGTTCGCTGCAACATGAGAAAGTGAAGCGTTACCTACACAGTGAACACCTG 8163
QY 3015 CGGCGTTTCAACATGAGCAACAGCGGCGAGTCCAGTATGAGAAATGGAATTAACGTA 3074
DB 8164 AAGGTTCTAACATGCACTGCAACAGCGGTGAGTGAAGGCAATTAATCAATCC 8223
QY 3075 CGAGAGAGTGGCGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3134
DB 8224 CCGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8283
QY 3135 TTGTGCTATTTGTTTGAAGAGTGCATATGAGGAGGAGGAGGAGGAGGAGGAGGAG 3194
DB 8284 TTGTCCGATGATGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8343
QY 3195 CTGGAACCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3251
DB 8344 CTTGAATGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8403
QY 3252 -----CACTAGTTACAGCGCTATGCGTCTTTCGATGTCAGTTCCTGCAACAA 3305
DB 8404 CTGACACCTGCTCAGCGCTATGCTTCTTCTGGAACGAGCTTCCATGCAATCGGC 8463
QY 3306 CACCGGTGCTATTTACTGACGCAAGCAACCACTGCACTGCTGCAAGAGAGAGTGC 3365
DB 8464 CGCCCATGCTACACCCCGGAACTATCCAGAGCTCTGCACTCTGCAAGAGAGAGTGA 8523
QY 3366 ACAATCAATTAACGACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3422
DB 8524 ACCAGAGGCTTACGACACCTGCTCAACGCAATTTGGGTCGGAATGCTCGGAGAA 8583
QY 3423 CGAAGAGGATTAACGATGACTTCACTGACAGTCTCTTCTGAGGAGTTCGCGCT 3482
DB 8584 GTAAAGAGGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8643
QY 3483 ATTGAGAGACTCAACGCGGTGTTTCAAGCCCAATTAATTAATGAGAACTGTGGAGCA 3542
DB 8644 ACTGTACATATGTAACCGTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8703
QY 3543 CTGATATGAGATGATTAATGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3602
DB 8704 CGAGAGCAACACCATACGATACAGATTCGCGCAAGTTGAGATACGACAAAGCGGAG 8763
QY 3603 CTGCGATGTCACCAATTCGTTACATGTCCTTTCAGCACGACGATGATCAATCAAGGAG 3662

Db 8764 CAGAAAGCTCAATAAGATACCGCTACATGCTCGACAGAGATATCTGTCAAGAG 8823
Qy 3663 ACAGATAGAGAAAAATAGCTATTCACACATCTGAGACCTGCGGTGTCTTGSCCAAG 3722
Db 8824 GCAACATAGATGACATCAAGATCAGACCTCAGACCGTGTAGAAAGCTTAGCTACAAAG 8883
Qy 3723 GGTATCTCTGTAGCTCAATGTCTCCAGGTGACAGTGTAAACCTGTAGTACAGACG 3782
Db 8884 GATACCTTCTCCTCGCAAGTGTCTCCAGGGGACAGGTAAACGTTAGCATAGAGATA 8943
Qy 3783 GAGCATCTGAGAAATTCATGACCGTGGAAAAAGATCAGAGGAATTTGTGGTAGAG 3842
Db 8944 GCAACTGACCAACGTCATGACCAATGGCCCGCAAGATATAACCAAAATTCGTGGACCGG 9003
Qy 3843 AGAGTACTTGTGCCACCCGCTCCATGGAAGCTGTAAAGTGCACGTTTACGATCCT 3902
Db 9004 AAAAATATGACTTACTCTCCGTTACGGTAAAGAAATCTTTCACATGTACACCGTCC 9063
Qy 3903 TGAAGAGACGCTCTGCGGGTACATTAACATGACAGGCCAGGCCACACCGCTATAAGT 3962
Db 9064 TGAAGAAACAAACCGCGGTACATCACTATGACAGGCCGGGACCGGACCGCTATACAT 9123
Qy 3963 CCTATCTGAGGAAACGCTCAGGCGAAGTGTACATTAAACACCTTGTGGCAAGACGTCA 4022
Db 9124 CCTATCTGAGGAAATCATCAGGGAAGTGTACGCGAACCAATCCGGGAAGAACATTA 9183
Qy 4023 CCTACGAAATGTAAGTGTGACCTACAGCAGAGTATGTGACAGCCGGGAACGAAGATGA 4082
Db 9184 CGTACGAGTGTCAAGTGGCGGCAATTACAGACCGGAACCGTTACGACCGGTACCGAAATCA 9243
Qy 4083 ACAGCTGTCACTAAAGCAAAACAGTGCATTCCTCAAGAGGAGCAACGAATGGGTCT 4142
Db 9244 CGGGCTGACCGGCATCAAGAGTGTCTGCTTATTAACAGCGACCAACGAATGGGTCT 9303
Qy 4143 TCAACTGCGCGGATCTTATTAAGGACACAGACCACTCAGTGCAGAGTAAATTCACATTC 4202
Db 9304 TCAACTGCGCGGATCTGATACAGACCGCGCAACACCGGCCAAAGGAAATTCGACTTTCG 9363
Qy 4203 CATTCGCTTGACACCGGACAGTGTGCGCGGTTCCTTACGACAGCGCTTACAGTACAGA 4262
Db 9364 CTTTACAGCTATCCCGAGTACCTGATGTCTCCGTTGCGCGCGCGCAACGATGATAC 9423
Qy 4263 AGTGTGTCAAAAGGATCACTCTCCACTGACCTGACATGACGACCAATTTGCTGACAAAG 4322
Db 9424 ACGGCTTTAAACATACATGACCTTCAATTAAGACAGACACATCTGACATTTGCTACACCA 9483
Qy 4323 GAAATATGGGCTGCGAGCAGACGCAACAGCAAGATTAACAGGCTCTACATCCAGGA 4382
Db 9484 GGAAGCTAGGGGCAAAACCGGAACCAACCACTGAATGATCATGGGAACACGGTTAGAA 9543
Qy 4383 ATTTTCTGTGGGGGAGAAAGGGCTGTGATGATGAGTAAAGGTTAACATACCACTACAG 4442
Db 9544 ACTTACCGTGTGACCGAAGATGCGCTGGAATCATATGGGGCAATTCACAAACAGTAAAGG 9603
Qy 4443 TCTGGGCGCCAGAGTGTGCGACAGGCGACCCACATGATGGCGGATGATCATATCC 4502
Db 9604 TCTATGCCCAAGATGTCTGACACAGAGACCTTACAGATGGGCTACACAAATAGTACAGC 9663
Qy 4503 ACTATATCATCGGATCGAGTCTACATGTATGTGCTGTGTGTGCTCTCTTGTGA 4562
Db 9664 ATTACTATCATCGCATCTGTGTACACATCTTAGCGGTGCGCATCAGCTGTGGGA 9723
Qy 4563 TCCGTGTAGGACATGATCATACAGAGTGTGATGCGCAAGAAAGAAAGATGCTCTGA 4622
Db 9724 TGAATATTTGGGTAACTGTGTGACAGATTAATGTGCTGTAAAGCGCGCTAGAGCTCTGA 9783
Qy 4623 CGCCATAGCGCTTGACACGAACGATACCAAGATTAAGCGGTTTGTGTGCA 4682
Db 9784 CGCCATATGCTTGGCGCCCAATGCGGTATTCACATCTTGGTACATTTGTGCTGTG 9843
Qy 4683 TTGGCGCAACCAAGCTGAAACATTTGGAAACCTTTGACCATCTGTGTTTAAACAAC 4742

Db 9844 TTAGTGGCTAATGCTGAAACATTCACCGAGACCATGATGATTAATGCTGAACAGCC 9903
Qy 4743 AACGTTTCTGTGGGACAGTTGTGATTCCTGTGGACGCGCTTGTATTTCTGTTCGCT 4802
Db 9904 AGCGTTTCTTGTGGGTCAAGTGTGTATCTTGTGGCGCTGTGTGTCTTTAATGCGCT 9963
Qy 4803 GCTTTTCAATGCTGATGCTTTTATTAATGTTGAGAGCGTGTGCTGGGAAAGTGAACG 4862
Db 9964 GTTGCTATGTGTGCTGCTCTTTTATGATGTGTGCGGGCTTACTGTGGCAAGTGAACG 10023
Qy 4863 CTTTGAACATGCAACATCTGTGCAAAATGTTCCGGGATTCGCTGTATAGCGCTGTGCG 4922
Db 10024 CTTAGCAACATGCAACATCTGTGCAAAATGTTCCGGGATTCGCTGTATAGCGCTGTGCG 10083
Qy 4923 AACGGAGGTTTGTGGGCACTTAACCTGGAGATCAGGTTGTCTATCGGAATTAACAC 4982
Db 10084 AAAGGAGGTTTGTGGGCACTTAACCTGGAGATCAGGTTGTCTATCGGAATTAACAC 10143
Qy 4983 CTTCACTAACAGAGTACGTGACCTGCAAAATTCACACAGTATTCCTTACACCAAG 5042
Db 10144 CTTCACTAACAGAGTACGTGACCTGCAAAATTCACACAGTATTCCTTACACCAAG 10203
Qy 5043 TTAATGCTGTGGGTCTCTGATGCAAGGCAATCTCAAGCGGATTAACATGCGCGG 5102
Db 10204 TCAGATGCTGTGGGTCTCTGATGCAAGGCAATCTCAAGCGGATTAACATGCGCGG 10263
Qy 5103 TTTTGGGCGGTGTACCTTTCATGTTGGGAGGCGGCAATGCTTCTGTGACATGAGA 5162
Db 10264 TCTTGGAGGGGTGTACCTTTCATGTTGGGAGGCGGCAATGCTTCTGTGACATGAGA 10323
Qy 5163 ACACAACTGAGTGTGAGGCGGTACGTGATTCCTGCAAGTGTGATTAATGATACAGCG 5222
Db 10324 ACAGCGAGTGTGAGGCGGTACGTGATTCCTGCAAGTGTGATTAATGATGACAGCG 10383
Qy 5223 TCGCACTAAAGTTCACACAGCTGCTGAAAGTGTGCGCTGTGTATAGTATGCGCAACA 5282
Db 10384 AGCGGATTAAGTGTGATTCGCGGATGAAAGTGTGATGATGATGATGATGATGATG 10443
Qy 5283 CCACCGGCGACCTGGATAGCTTGTGCAATGCGGTGACCGGACGTTCTCACGGGACCTGA 5342
Db 10444 CTACCACTTCTCTGATGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 10503
Qy 5343 AGCTCATAGCAGGCGCATATACAGCCCTTTTTCACCTTTTGAACATTAAGTGTGATCA 5402
Db 10504 AAGTCATAGCTGAGCAATTCATGATTTTGAACATTAAGTGTGATGATGATGATGAT 10563
Qy 5403 GAAAGGCGCTTGTTAACACTACGACTTCTCTGATGATGATGATGATGATGATGATGAT 5462
Db 10564 ATCGCGGCTGTGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 10623
Qy 5463 TCGGCAATTTCAAGCATCTCTGCTTATGATGATGATGATGATGATGATGATGATGATG 5522
Db 10624 TTGGAACATTTCAAGCATCTCTGCTTATGATGATGATGATGATGATGATGATGATG 10683
Qy 5523 GCGTGTGAAGCTTCTGTCAAGAAATTCACAGTCCCTTACACCAAGCAATGATCAGG 5582
Db 10684 GCGTGTGAAGCTTCTGTCAAGAAATTCACAGTCCCTTACACCAAGCAATGATCAGG 10743
Qy 5583 ATGAATGTGGAAGAAACATCTCAGAGACCCCTGCAAGAAACAGCAATTTGATGATGA 5642
Db 10744 TCGAGATGTGGAAGAAACATCTCAGAGACCCCTGCAAGAAACAGCAATTTGATGATGA 10803
Qy 5643 AAATTTGAAGTGAAGCTCTGTGAGAGCTTAACTGTGTTTACGGGCAATCTCTATCTGGA 5702
Db 10804 AGATTTGAAGTGAAGCTCTGTGAGAGCTTAACTGTGTTTACGGGCAATCTCTATCTGGA 10863
Qy 5703 TTGACATCTGATGAGAGCTTTGTGATGATCATCAGATCAACCAATTTTGAAGTTA 5762
Db 10864 TTGACATCTGATGAGAGCTTTGTGATGATCATCAGATCAACCAATTTTGAAGTTA 10923
Qy 5763 GCTGACAGTGAAGAGCTGATTAATTTCTGACAGCTTTGTGATGATCATCAGATCA 5822
Db 10924 AATGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10983

[illegible]

D	5827	AAAGAAATCTAGCCCGCGTGTCTGCACACGTCTGAAAGAGAAAGCAAGCTCAAACTCAGTAACC	5886
Q	760	AATTAATGCGCTCTGAAGAAATATGAAAGCAGGTATCAATACGAAAAGTAAATATATGA	819
D	5887	AGATGATGCCACCGAAGCCAAACAAAGACAGGTACCAGTCTCGAAAAGTGAACGAGA	5946
Q	820	AAGCAATTACAGCGGACGACCTCATTTCTGATTTGGCACATATCTATCTCATCAGAACTGA	879
D	5947	AAGCCATTAACCACTGACGACCTGCTTTTCAGGGCTACGCGTGTATTAACCTCTGCGCACAG--	6003
Q	880	ATCTGTTCGAGTGTTAACAGAGTCAATATTCCTGTACCAATCTACTCGTCAACGGTAAATTA	939
D	6004	ATCAGCCAGATATGTTATTAAGATCACTTACCCGAAGCAATCGTATTTCCAGCAGTGTACAG	6063
Q	940	ACAGGTTTACATCTGACGAGGTCGCGGTAAACGTGCACACTTAAGTTATCCAGAGAAATT	999
D	6064	CGAACTACTCTGACCCCAAGATTTGCTGTAGCTGTTTGAACAATCTATCTGCATGAGAAATT	6123
Q	1000	ACCCTTACAGTATGCCAGTTATTTGTTATTAACAGATGAATACGATGCTATCTTGACATGGTGG	1059
D	6124	ACCCGACGATGACATCTTATTCAGATCACCGACGATGACGATCTTACTTGGATATGGTAG	6183
Q	1060	ACGGCGCATGTCGTCTATGATACACACCACTTTTGTCCGGCTTAAACTGAGAGCTATCC	1119
D	6184	ACGGGACAGTCTGCTTCCCTAGATATCTGCAACTTTTGTCCCGCCAGCTTATGAAATTAACC	6243
Q	1120	CAAGAAGCATAGCTATTTTGGACGCCAGAGATTAAGATCAGCGCTCCATCGCTATACAGA	1179
D	6244	CGAAAAGCACAGATATTAAGACCCCAACATCCGACAGTGGGGTTCCATCAGCCAGATGACGA	6303
Q	1180	ATACATTACAAATATGTAATTTGGCTGCACTACTTAAAGAAATTCGAACGTTATCCCAATATGC	1239
D	6304	ACAGGTGGCAAAAGTGTCTCATTTGCCCGCACTAAACAACTGCAAGTGTACACAAATATGC	6363
Q	1240	GAGAAATTAACCTGTCTTATGATTCGGCGGGCATTTAATGTTGATGTTTCAAGAAATTAACCAT	1299
D	6364	GTAATCTGCCCAACTGGAATCTGACGACGATTTCAAGTTGATGCTTTGCAAAATATGATAT	6423
Q	1300	GCAATGATGATCTACTGGGATACCTTTTCGCGATTAACCTTATTCGGCTTAACTACAGAGAACG	1359
D	6424	GCAATGACAGATATTTGGGAGGAGATTTGCCGAAAGCCAAATTAGGATCACTACTGATTTGC	6483
Q	1360	TTAGCGCAATATGTACACAAAGCTGAAGGGCCGAAAGACAGACGATTTGTTGGCAATATCTC	1419
D	6484	TTACCGCATACGTGTGGCCAGACTGAAGAGCCCTTAAGGCCGCCGCACTTTTCGCAAAAGACGC	6543
Q	1420	ATTAATCTTAAACCGTGTGAGGAGATATCAATGATCAATTCGTATGATGATCTTAAAGAGAG	1479
D	6544	ATTAATTTGTCTCCATTGCGAAGATGCTGTATGATTAAGTTATGTTATGATGACATGAAAAAGAG	6603
Q	1480	ATGTCAAAAGTTATCTCCGGCCAGCAAAACATACAGAGAGCGGCTTAAGGTGACAGTTATTC	1539
D	6604	ACGTGAAGATTAACACCTGTGACGAAACACACAGAAAGAAAGCCGAAAGTTCAATGATATAC	6663
Q	1540	AGGCTGCAGATCCCTTGTCTACCGCTTACCTTTGCGGGATTCATCGGGAAATTAGTCCGTA	1599
D	6664	AAGCCGACGAAACCCCTGGCGACCCGCTTATCTATTCGGGATATCAACCGGAGATTTATGTGCGCA	6723
Q	1600	GACTGAATACGCGTCTCTCTGCCAAATATACATACTCTCTTGCAGATGTCAAGCGGAAGATT	1659
D	6724	GGCTTACAGCCGTTTGTCTGATCCCAACATTCACAGCTCTTTTGAATATGTGCGGAGAGACT	6783
Q	1660	TTGATGCGATTAATTTGCTGAACATTTTCACACGCGGACCCAGTATTTGGAAACGACATATCG	1719
D	6784	TTGATGCAATATATGACAGAAACATTTCAAGCAGGTGACCCGATATCTGAGAGCGGATATTCG	6843
Q	1720	CGTCGTTGATAAAGCGAAGACGACGCTATCGCATTTTCGCGCTGATGATCTTTGAGG	1779
D	6844	CCTGCTTTCGACAAAGCCAAAGACGACGCTATGAGCGTTTAAACGCGGCTGATGATCTTTGGAAG	6903
Q	1780	ACTTAAGTGTGACCAACCGCTCTTAATATTTGATATTAAGAGCGCGCTTTCGGCAATATCATAT	1839
D	6904	ACCTGGGTGTGACCAACCACTACTGCACTTGATATGAGTGTGCGCTTTTGGAGAAATATCAT	6963

1840 CTGTGACCTTACCTACAGAAAGAGTTTAAATTTGTGTCATGTAATCCGTAATG 1899
6964 CCACCACTGTGCGCCAGGGATACCGTTTCAATTCGGGGCATGATGAATCCGAAATGT 7023
1900 TCTTAAAGCTGTGTTGTCAACACACTATGATATATATGATGTCAGAGTAATCGTG 1959
7024 TCTTCAAGCTCTTTTGTCAACACAGTTTGAATGTGTTATCGCCAGCAGATTTGAGAG 7083
1960 AACGCTTAAACACGTCAGCGTGGCGGCTCTATGCGGACATTAACATAGTCATGTG 2019
7084 AGCGCTTAAACGTCCAAATGTGACAGATTTATGGGACGACAACTTATATACAGGAG 7143
2020 TCGTCTCCGACACTTGTGAGAGAGATGCGCACTTGGCTGAAACATGAGTAATAA 2079
7144 TAGTATCTGACAAATAATGCTGAGAGGTGTGCCACTGCGTCAACATGAGAGTTAAGA 7203
2080 TTATGATGACGTTATTTGGTATCAAGACCTTACTTGTGTGGGGATTTATCTGTGTG 2139
7204 TCATTTGACGAGTCATCGGCGGACAGACCTTACTTGTGCGGTGATTCATCTTGCAAG 7263
2140 ACCGATTAACAGGCGACGCTGCAGATGCGAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
7264 ATTGGTTAATCTCCACAGCGGTGTGCGTGGCGGACCTTGAAGGCTGTTTAAAGTTGG 7323
2200 GAAAACCAATGCGCACTGATGATACCAAGACTGCGACCGCGCGGCACTGATGATG 2259
7324 GTAAACCGCTCCGACCCACATGAGCAAGACGAAAGACAGAAAGCGCTGTCTGTAAGTG 7383
2260 AAGCAATCGATGGAACAGAAATTTGAAATTAACGAGATGTAAGGCGGTGAATCCA 2319
7384 AAACAAAGGCGGTGTTAGAGTATTAACAGACCTTGAAGCAATGAGCGGTGCAATC 7443
2320 GATACGATCATATCTGGCAGGCTGATCATCACTCTGTGCCAGTTAGCCGAAAGCG 2379
7444 GGTATGAGTATGACAAATCAACACTGTCTGTGCAATGAGAACTTTTGGCCGAGACA 7503
2380 TTAAGAACCTTCAAGACATTAAGAGGAGCGCAATCACTCTAAGGCTGACCTTAAATAG 2439
7504 AAAGAGCACTTTCAGACCATCAAGAGGAAATTAAGCATCTCTACGAGTGTCTTAATAGT 7563
2440 TGACGTACTAGA-----CAGCACTTACCCACCGGAGATGTTTCCA 2482
7564 CAGCATATGATCTTCTGATCTGACTATATACACACACACCATGATGAGATTTCT 7623
2483 TACCTCAGCTGAATTTTCCACCAATTAACCTTAACAATCGATGCTTACCGAGATCCA 2542
7624 TTAAATGCTCTGGCGCGCGCCCTTCCAGCCCCCACTGCAATGTGAGGCGCGAGAA 7683
2543 AACCTCTTAGGCGCGCTGAGGCGGTTTGGGCCCGCTGCTGCTCAATGCAAT 2602
7684 GAGGACAGGCGGCCCGATGCTGTGCCGCAATGGCTGCTTCCAAATTCGACAACTGA 7743
2603 CTTAGAGAGTGAATGATCACTTGAATTTCAAAACAAAC--GATCACTTAATCGCGCCAG 2660
7744 CCACAGCGGTGATGCTCTGATGATGACAGCAATAGACCTTAATCCCAACGCCAC 7803
2661 GTCCACCGCCCAAGAAAGAAAGAGTGTCTTAAGCCAAACCTAATGAGCTTAATAAGA 2720
7804 GCGCGCCCGCGCGCAGAAAGAGGCGCCAAACCAACCGAAGCGAAGAAACCA 7863
2721 AGAAGCAGCAAGCCAAAGGAGCGAAGAAC--GCAAGCTTAAACAGAGAAAGCAACCTA 2777
7864 AAACACAGAGAAAGAAAGAAAGCAACTGCAAAACCCAAACCGGAAAGAGACGCTGA 7923
2778 TGTGTATGAGTTGAGTCTGACAAAGATTTCCGATC--ATGCTGAACGCGCAAGTGA 2834
7924 TGGCACTTAATGTTGAGGCGGACAGACTGTTGAGAGTCAAAAATAGAGACGAGATGCA 7983
2835 ATGATATGCTGCTGTGTGAGAGAGGCTGATTAACCACTCCAAGTTGAGAGAAA 2894
7984 TCGGCGACGCACTGCGCATGAGAAAGAGTATGATAACCACTCCACGTAAGAAAGAACTA 8043

2895 TTGATATAGCAATTACCGCGCGTGAATTTGAAGAGGCTAGCATGACACTTGGAGT 2954
8044 TTGACCAACCCGTGTGTATCAAGCTCAATTCACCAATCTGTACGATACGACATGAGT 8103
2955 ACGCGACGTTTCCCGAAGCATGAATCAGACAGCTGCAGTACACAGCGCAAAACAC 3014
8104 TCGCACAGTGTCCCGTCAACATGAGAGTGAAGGCGTTTCACTTACACAGTGAACCTTG 8163
3015 CGGCTCTTACAACTGGGACCAACCGCGGAGTCCAGTATGAGAAATGGGATTTTACCGTAC 3074
8164 AAGGTTCTTAACTGGGACCAACGAGGCGGTGAGTATGAGAGGAGATTTTACATTC 8223
3075 CGAGAGAGTGGGCGGAAAGGCGACAGCGGAAGCCGATCCTGACAAACAGAGCAAG 3134
8224 CCGCGGAGTATGAGAGGACAGAGAGACAGTGTGTGCTGCAATATGATTAATCAGGCGGG 8283
3135 TTGTGGCTATTTGTTAGAGAGTGAATGAGGCGCGCTTACGCGCTTTTCAAGTGTCA 3194
8284 TTGTGCGATATGTCCTCGAGGGGCGTGAATGAGGAAACGAGCGCTTTCGCTGTCA 8343
3195 CTTGGAAACAGAAAGGGGTGACATTTAGGATTAACCCCGAAGGTTCTGAACCGTGT--- 3251
8344 CTGGATATGCAAAAGGAGAAACAATCAAGACACCCCGAAGGACAGAAAGTGTGTG 8403
3252 -----CACTAGTTACAGCGCTATGCGTGTGTTGCAATGTACGTTCCCATGCGCAAAAC 3305
8404 CTGACACCTGTGTACGCGCATGTGTCTTGTGAAAGTGAAGTCTTCCATGCAATGCC 8463
3306 CACCCTGTGTATTTACTGACGCGAGACGCAAGCACTGACGTGTGAGAGAAAGCTGCG 3365
8464 CGCCCACTGTCTTACCGCGCAACCAATCAGAGCTCTGACATCTCTGAAGAGAACTGA 8523
3366 ACAATCAAAATTAAGACACGCTGCTGAGAAAGCTTGAAT--GTCAATACGCGCGC 3422
8524 ACCAGAGGCTTACGACACCTGTCTCAACGCAATATGCGGTGGATGTGTCGCGCAAA 8583
3423 CCAAGCAAGGATTAACGATGACTTACATGACAGCGCCCTTACCTGGGGTTTGTGCCGT 3482
8584 GTAAAGAGGCTCACTGACATCTTACCTTGAACGAGCGCTTACCTTGGGACATGTCTGT 8643
3483 ATTGCAGACTCAACGCGGTGTTTCAAGCCCAATTAATTAATGAGAACGTGTGGAGCAAT 3542
8644 ACTGTACCATATGAAACCGTGTGCTTGAAGCCGATTAAGATGAGAGAGTCTGGAGTAA 8703
3543 CTGATGATGATTCGATTAAGATTCAGAGTCTGCGCAATTTGGCTTACATGAGGAGCA 3602
8704 CGAGACAAACCAATAGCATACAGACTTCCGCGCAAGTTGATGATGACCAAAAGCGGAG 8763
3603 CTGCGGATGTACCAAAATTCGTTTACATGTCTTTCGACACAGACCATGACATCAAGAG 3662
8764 CAGCAAGCTCAAAATTAAGTACCGCTTACATGTCTGCTGAGACAGATCTATCTGTCAAGAG 8823
3663 ACAGATGAGAGAAATAGCTATGACACATCTGACACCTGCGTGTCTTGGCCCAAG 3722
8824 GCACATGATGATGATCAAGATGACAGCACTCAGAGCGGTGATGAAGGCTTATGATCAAG 8883
3723 GGTACTTCTTATGCTCAATGTCTTCAAGGTGACAGTGTAAACGCTCATGTATCAGAGCG 3782
8884 GATACTTCTCTCGGAAAGTGTCTTCCAGGGGACAGGTATTAACGTTAGCATAGGAGTA 8943
3783 GAGCATCTGAATTAATGACACCGTGGAGAAAGATGAGAGAGGATTTGTGGTAGAG 3842
8944 GCAACTGACGACGTCATGACCAATGAGCGCGCAAGATTAACCAAAATTCGTGGAGCGG 9003
3843 AGGATTAATGTTTCCACCCGTCATGAAAGCTGTAAATGTGACACGTTTACATCACT 3902
9004 AAAAATATGACTTACTCTCCGTTTACCGGTAAAGATTCCTTGACAGTGTACGCGCTC 9063
3903 TGAAGAGAGCTCTGCGGGTATCAATTAACATGACAGGCGCCACAGCGCTTAAAGT 3962
9064 TGAAGAAACAAACCGCGGCTTACATCTTATGACAGGCGCGGACCGCATCTATCAT 9123
3963 CTTATCTGAGAGAGCGTCAAGCGAGTGTATTAACCACTTGTGGCAAGAACGTC 4022

```
Db 9124 CCTATCGAGAGATTCATCAGGAAAGTTTACGGAGACCATCCGGAGAAACAATTA 9183
Qy 4023 CCTACGAATGTAAGTGTGGCACTACAGACAGATATGTGACACCGGAAAGAAATGA 4082
Db 9184 CGTACGAGTGAAGTGTGGCACTTACAGACCGGAAACGTTTACACCCGTCCCAAAATCA 9243
Qy 4083 ACGGCTGCATTAAGCAAAACAGTGCATTGCTTCAAGAGGACCAAAAGAAATGGGCT 4142
Db 9244 CGGGCTGCACCGCCATCAGAGATGCTGCTTATAGAGGACCAAAAGAAATGGGCT 9303
Qy 4143 TCAACTGCGCGGATCTTATAGGCAACAGACCACTAGTCAAGTAAATTTGACATTC 4202
Db 9304 TCAACTGCGCGGATCTGATCAGACACCGGCAACAGCGGCAAAAGGAAATTTGATTTGC 9363
Qy 4203 CATTCGCTTACACCGACAGTCTGCCGGTTCGTTAGCTCAGCCGCTTACAGTACGA 4262
Db 9364 CTTCAGCTGATCCCGATCTGACATGATGCTCCGTTGCCCACGGCCGGAACGTATAC 9423
Qy 4263 AGTGTTCAAAGGATCACTCTCACCTGACCTGCAATGCAACCAATTTGCTGACAA 4322
Db 9424 ACGGCTTTAAACATCAGCTTCAATTTAGACAGACCATCTGACATTTGCTTACACCA 9483
Qy 4323 GAAATTTGGGCTCGAGCAGACCAACAGAGATTTACAGGCTTACATCCAGA 4382
Db 9484 GGAAGCTAGGGGCAAAACCGGAAACCACTGATGATCATGGAAACAGGTTAGAA 9543
Qy 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTACGTATGGGGTAAACATTAACCACTAGAG 4442
Db 9544 ACTTCACGCTGACCGAGATGGCTTGAAATCATATGGGGCAATCACCAACAGTTAAGG 9603
Qy 4443 TCTGGGCGCAGAGTCGGCACCGAGCGACCAACATGATGGCGGATGATCATATCC 4502
Db 9604 TCTATGCCCAAGATCTGACACAGAGACCTTCAGAGATGGCCACAAATGATGACGC 9663
Qy 4503 ACTATATCATCGGATCAGCTCACTGTCATTTGCTGTGTGTGCTCTTGTGTA 4562
Db 9664 ATTACTATCATCGCATCTGTGTACACATCTTAAGCGTCGACATGCTGTGGGGA 9723
Qy 4563 TCTGTGAGGACATGATCATGACAGCTTGCATGCGCAAGCAAGAAAGATGCTGTA 4622
Db 9724 TGATGATTTGGGCTAATGTTGACAGATTTATGCTGTAAAGCGCGCTGATAGTCTGTA 9783
Qy 4623 CGCCATACGGCTTGACCGAAGCAAGGTAACCAAGCATTTAGCGGTTTGTGCTCA 4682
Db 9784 CGCCATATGCTTGGCCCAAAATGCGTATTCAACTTCGCTGACATTTTGTGCTG 9843
Qy 4683 TTCGGCAACCAACGCTGAAACATTTGGAGAAACCTTGTGACATCTGTGTTTAAACAC 4742
Db 9844 TTAGTGGGCTAATGCTGAAACATTCACCGAGACATGATTAATTATGTCGAACAGCC 9903
Qy 4743 AACCGTTTCTGTGGGACAGTTGTGATCTCTGTGGACGCGCTGTTATTTCTGTTCCGCT 4802
Db 9904 AGCGGTTCTTGTGGGCTAGCTGTGTATACCTGTGGCGCTGTGTCTTTTAAAGCGCT 9963
Qy 4803 GCTTTTCACTGATGCTTTTATGTTTATGTTGACGCGTCTGCTGGGGAAGTGAAGC 4862
Db 9964 GTTCTCATGCTGCTGCTTTTATGTTGTTGCGGCGCTTACTGTGCGAAAGTGAAGC 10023
Qy 4863 CTTTTCGACATGACACACTGTGCCAAATGTTCCGGGATCCCGTATTAAGCGCTGTGCG 4922
Db 10024 CTTTTCGACATGACACACTGTGCCAAATGTTCCAAATGTGCCAAGATACCGTATTAAGCACTGTG 10083
Qy 4923 AACGCGAGTTTACGCGGCACTTAACTGTGAGATCAGGTGTCTCATGCGAATTAACAC 4982
Db 10084 AAAAGGCGAGGTAGCGCCCGCTCAATTTGAGATTAATCTGTATGCTCTCGAGGTTTTC 10143
Qy 4983 CTTCAACATAAGAGATAGTACCTGCAATTCACAGATTAATTCCTTACACAGAG 5042
Db 10144 CTTCAACATAAGAGATTAATCTGCAATTAATTCACAGATGATGCTCTCTCCCTTAAAG 10203
Qy 5043 TTAATGCTGGGCTCTCGAGTGAAGGATCCTCAAGGCGATTAACATGCGCG 5102
Db 10204 TCAGATCTGCGGCTCTTGAATATGTCAGCCCGCTCAGCAGACATATACCTGCAAG 10263
Qy 5103 TTTTGGCGGTGTGTACCTTTTCAATGTTGGAGGCGCAATGCTTGTGACAGTGA 5162
Db 10264 TCTTTGAGGAGGTGTACCTTTTCAATGTTGGAGAGAGCAATGTTTGTGACAGTGA 10323
Qy 5163 ACACACAACTGAGTGAAGCGCTAGCTGAGTTCGCTCCAGACTGCATATAGTACAGCAG 5222
Db 10324 ACAGCCAGATGAGTGAAGCGCTAGCTGATTTGTATGATGATGTCGACATGACAGCGC 10383
Qy 5223 TCGCATTAAGTTTACACAGCTGCTTGAAGTGGCTGCTGCTATATGATAGCGCAACA 5282
Db 10384 AGGCGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 10443
Qy 5283 CCACCGGCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5342
Db 10444 CTACAGTTCCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 10503
Qy 5343 AGGTCATGAGAGGCGCATATCAGCCGCTTTTCAACCTTTTGAACATAGGTGCTCATCA 5402
Db 10504 AAGTCATAGCTGAGACCAATTTTCAAGCATTTTATACACATTCATGATCAAGAGTGTATCA 10563
Qy 5403 GAAAGGCGCTGTTTATACATGACATTCCTCTGATATGAGACTATGAACAGAGGCT 5462
Db 10564 ATCGCGGCTGTGTGTACATGATGATTCGGAATAGGAGGCGATGAACAGAGGCT 10623
Qy 5463 TCGGCGATTTCAAGCATTCCTGCTGATGCTCAGCATATGATGCGGACTGACATAC 5522
Db 10624 TTGAGACATTTCAAGCATTCCTGCTGATGCTCAGCATATGATGCGGACTGACATATTA 10683
Qy 5523 GGTGCTGAGAGCTTCTGTCAAGAAATCCACGCTCCCTTACACCAAGCATATCAGGCT 5582
Db 10684 GGTACTACAGCTTCCGCAAGAGCGTATGCTCCGTAACAGCAAGCGGCACTGAT 10743
Qy 5583 ATGAATGTGAAGAACATCTCAGAGACCCCTGCAAGAAACAGCAATTTGATGTA 5642
Db 10744 TCGAGATGTGAAGAAACATCTCAGAGCGGCGGCTGCAAGAAACCGGCTTTTGGGTCGA 10803
Qy 5643 AAATGAAGTGAAGCTCTGCGAGCGCTTACATGCTGCTTACGGGCAATCCATCTGCA 5702
Db 10804 AGATGAGTGAAGCTCTGCGAGCGGCTTACATGCTGCTTACGGGCAATCCATCTGCA 10863
Qy 5703 TTAGCATCTCCGATGACAGCTTTTGTGATCATCAGATCAACCAATTTTGAAGTGA 5762
Db 10864 TTGACATCTCCGAGAGCGCTTATATCAGACATCAGATGACACATGCTGTCAACATCA 10923
Qy 5763 GCTGCAAGTGAAGCTGATTAATTTCTGCAAGCTTTGTGCTTCTTAACTTACAT 5822
Db 10924 AATGTGATGTCAGTGAAGTGAAGCTTATCAGCGGCTTGGAGGATGCTTACCTGCAAT 10983
Qy 5823 ACAAGCTGACAGGAGAGGACATTTGTCAGTTCACCTCCACTCCAGACAGCTGTTTGA 5882
Db 10984 ATGTATCGAGCGGAGAGCAATGCTCTGTACATTTGCAATTTGAGCAGACCAACCTCC 11043
Qy 5883 AGAAGGCGACACATGATGATGCTCCGTAGCAGACATTAACATTAATTTAGACATGCA 5942
Db 11044 AAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 11103
Qy 5943 GCCCACAAGCAATTTATATGTTTCTGATGCGGCAAGATCCACTGCAATGCTGAAT 6002
Db 11104 GCCCACAAGCAATTTATGTTTCTGATGCGGCAAGATCCACTGCAATGCTGAAT 11163
Qy 6003 GTAAACACAGCGGCGACACATTAATGGAACCACTAAAGTGAACCAAGAAATTCAGG 6062
Db 11164 GCAACCAACAGCGTGAATATGTTGAGCAACCCGCAACCAAAATGACCAAGAAATTCAGG 11223
Qy 6063 CGGCAATTTCAAAACATCTTTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6122
Db 11224 CCGCATCTCAAAACATCTTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11283
Qy 6123 TCAATGTTGAGCACTTATATGTTGCTGAGCTGATGCTTATTAACACAGTATGAT 6182
Db 11284 TATTAATTAAGGCTTATGATTTTGTCTGACAGATGCTGATGCTGATGCTGATGCTGATGCT 11343
```

QY	6183	GACTGACCGCGGACACTGA-CATGCGCGTAAACTCGANTACTTCGCGAGGAGCGTGT	6241
Db	11344	GACGGCTACGCCCAATGACATCCGACCGACCAAACTCGATGACTTCGAGAACTGATGT	11403
QY	6242	GCATTAATGCCAGCGCGCGCT-----TGACACT	6268
Db	11404	GCATTAATGCATCAGGCTGTGTATATTAGATCCCGCTTACCGCGGCGCAATATACCAACCC	11465
QY	6269	AAACTGAGTATTTTCCGAGGAAGCACAGTGCATTAATGCTGTGCAGTGT	6318
Db	11464	AAAACCTGAGTATTTCCGAGGAAGCGCCAGTGCATTAATGCTGTGCAGTGT	11513
RESULT 7			
LOCUS	BD071122	11663 bp	DNA linear
DEFINITION	System for the in vivo delivery and expression of heterologous genes in the bone marrow.		
ACCESSION	BD071122		
VERSION	BD071122.1	GI:22616725	
KEYWORDS	JP 2001515348-A/3.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 11663)		
AUTHORS	Johnston,R.E., Davis,N.L. and Simpson,D.A.		
TITLE	System for the in vivo delivery and expression of heterologous genes in the bone marrow		
JOURNAL	Patent: JP 2001515348-A 3 18-SEP-2001; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL		
COMMENT	OS Girdwood S.A. PN JP 2001515348-A/3 PD 18-SEP-2001 PF 18-FEB-1998 JP 1998536760 PR 19-FEB-1997 US 08/801263 PI ROBERT E JOHNSTON,NANCY L DAVIS,DENNIS A SIMPSON PC CI:2N15/86,CI:2N15/33,CI:2N7/01,CI:2N5/10,A61K39/12,A61K48/00 CC System for the in vivo delivery and expression of heterologous genes in CC the bone marrow CC Key Location/Qualifiers FH 1. 11663 FT source /organism='Girdwood S.A.'		
FEATURES	source 1. 11663 Location/Qualifiers /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"		
BASE COUNT	3289 a	3105 c	2910 g 2359 t
ORIGIN			
Query Match	37.0%	Score 2397.6;	DB 6; Length 11663;
Best Local Similarity	64.6%	Pred. No. 0;	
Matches 3752; Conservative	0;	Mismatches 1984;	Indels 74; Gaps 9;
QY	580	ACTGACCGTATGAAAGCGGAGCGTATATTTCTCATCGGAAACAGGCGCAAGTCACTTC	639
Db	5707	ACTGTCTAACCGGGGTAGTGGTGTACTATATTTTCAGAGCACAGGCGCTCGGCACTTGC	5766
QY	640	AACGAAATAGTACGTCAATGTAACTACAGAACCTTATATTGATCGGCGCGTCCATG	699
Db	5767	AAAAGAAGTCCGTTTCTTCGAAACCAAGCTTACAGAACCGACCTTGAGCGCAATGTTCTGG	5826
QY	700	AGAAGTATTACGCCCGCGCGCTCGATCTCGAAGAGAGAAAAGTTACAGAAAGAACTGC	759
Db	5827	AAAGAACTTACGCCCGCGCGCTCGAAGAGAGAAAAGTTACAGAAAGAACTGC	5886
QY	760	AATTATGCGGCTCTGAGAAATAGAAAGCAGTATCATACGAAAGTAGAAAATATGA	819
Db	5887	AGATGATGCCCAACGAAAGCAAAAGACAGTACAGTTCGAAAGTAGAAAACACAGA	5946
QY	820	AAGCAATTACAGCGGAGCGCACTATTTCTGATTTGGGACATATCTATCATCAGAAATGA	879

Db	5947	AAGCCATPACCACTGAGGACCTGCTTTACGAGGCTACAGGCTGTATTAACCTCTGCCACAG---	6001
Oy	880	ATCCTGTCGAGTGTACAGAGTCAATTAATCTGTACCAATCTACTCGTCAACGGTAATTA	939
Db	6004	ATCAGCCGAAATGTCTATATGATCACTACCCGAAACCATGTGATTTCCAGCAGTGTACAG	6061
Oy	940	ACAGATTACATCTGCAAGAGTCCGGTTAAACGTGCACCTTACTTATCCAGAAATTT	999
Db	6064	CGAATCTCTGACCCCAAAAGTTTCTGTAGCTGTTTGTATCAACTATCTGATGAGAAATTT	6122
Oy	1000	ACCCATACAGTACCGAGTATATGTATTAACAGATGATATACATGCGCATCTTACATGTGAG	1051
Db	6124	ACCCGACGGTATGCACTTTATCAGATCACCGACGATGAGATGCTTACTTGATATGTGATG	6181
Oy	1060	ACGGCGATCTGCTGTCTAGATACAGCACTTTTGTCCGCTAAACTGAAAGCTAC	1111
Db	6184	ACGGGACAGTGCCTGCTGCTATATCTGCACTTTTGGCCCGCCAGCTTGAAGATTAC	6241
Oy	1120	CAAAAGACATAGCTATTTGCAAGCCAGAGATAGATACGCCGTCCATCGCCTATACGA	1171
Db	6244	CGAAAAGCACAGATATAGAGCCCCAAACATCCCAAGTGCAGTTCACTACAGATGCAGA	6301
Oy	1180	ATATCATTAACAAATATGTATTTGCTGACGTACTATAAGAAATTTGCAACCTTACCCAAATGC	1231
Db	6304	ACAGCTTGCAAAAGTGTCTCATTTCCCGCATTAAGAAATGTCAACCTCACACAAATGC	6361
Oy	1240	GAGAAATTAACCTGTCTTATGATTCGGCGGCACTTAATGTTGATTTGTTTCAAGAAATACCAT	1291
Db	6364	GTGAATCTGCCAAGCACTGGACTCAGCGACATTTCAACGTTGATGCTTCGAAATATATGAT	6421
Oy	1300	GCAATGATGAGTACTGGAGATACCTTTGGCGATAACCTTATTCGCTTAACTTACAGAAATCG	1351
Db	6424	GCAATGACGAGTATTTGGGAGGAGATTTTCCGAAAGCCAAATTAGGATCACTACTAGTTTCG	6481
Oy	1360	TTAGCGAATATGTACAAAGCTGAAAGGCGCGAAAGCAGACGATGTTTCCGAATATCT	1411
Db	6484	TTACCGCATAGTGTGGCCAGACTGAAGGCCCTTAAGCCGCGCACTGTTCCAAAGAGCG	6541
Oy	1420	ATTAATCTAAAAACGTTTSCAGAGAGATACCAATGATCAATTTGTCATGGATCTTAAAGAGAG	1471
Db	6544	ATATTTGTGTCCCTTGTGCAAGAAATGCTCTATGATATGTTGTATGTACATGTAAAGAG	6601
Oy	1480	ATGTCAAAAGTTACTCCGGGCAAGAAACATACAGAGAGCGGCTTAAGTGCAGTTATTC	1531
Db	6604	ACGTAAAGTTATCACTGTGCAAGAAACACACAGAAAGAACCGAAAGTACAAAGTATAC	6661
Oy	1540	AGGCTGCAAGATCCCTTGTGCTACCGCTTACCTTTTGGGATTCATCGGGAATTAATCTGCTA	1591
Db	6664	AAGCCGACGAAGCCCTGTGGCGACCGCTTACCTATCCGGGATTCACCGGAGCTTAATGTGCCCA	6721
Oy	1600	GACTGAATGGGGTCTTCTGCGAAATATACATATCTCTTCGACATGTCAACGGGAAGTT	1651
Db	6724	GGCTTACAGCGTTTCTACTCCAAACATTCACAGCTCTTTTGACATGTGCGCGAGAGCT	6781
Oy	1660	TTGATGCGATTAATTTGCTGAACATTTTCCACACGCGGACCCAGTATTTGGAAACGACATCG	1711
Db	6784	TTGATGCAATCATATGACGAACAATTTCAAGCAAGTGAACCGGTACTGGAAGCGGATATCG	6841
Oy	1720	CGTGTGTTGATTAAGAGGAAAGCAGACGCTATCGCAATTTCCGCGTTGATGATCTTTAGG	1771
Db	6844	CTCTGTTCGACAAAGCCAAAGCAGACCGCTATGCGGTATACCGGCTGATGATCTTTGGAAG	6901
Oy	1780	ACTTAAAGTTCGACCAACCGCTCTTATATTTGATATGATAGAGGGCGGCTTGGCAATATCAAT	1831
Db	6904	ACCTGGATGTGACAAACCACTACTGCACTTGATCGAGTGGCGCTTTGAGAAATATATAT	6961
Oy	1840	CTGTGCACTTAACCTTACAGAAACAGATTTTAAATTTGTGTGCATGATGAAATCCCGTATGT	1891
Db	6964	CAACCCATCTGTGCCAGGGTATCCGTTTCAAAATTTCCGGGGGATGTATGAATCCGGAATGT	7021
Oy	1900	TCTTAAGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGATACTTACGTG	1951

D 7024 TCCTCAGCCTCTTGTCAACAGCACTTCTGAATGCTTATTCGCCAGAGATATTGGAG 7083
Q 1960 AAGCGTTAACCAAGCTCAGCGGTGGCGGCTTATCGCGAGATAGATAGTGAATGTG 2019
D 7084 AGCGGCTTAAACCTCAAAATGTGACATTTATTCGGGAGAGCAAACTTATACAGAG 7143
Q 2020 TCGTCTCCGACACCTTGAATGGCGGAGATGCGCCACTTGGCTGAACATGGAAATTA 2079
D 7144 TAGATCTGACAAAGAAATGGCTGAGAGTGTGCTACCTGGCTCAACATGAGGTTAA 7203
Q 2080 TTATTTAGATGAGTATTGTTGATCAAGACCTTACTTGTGGGGATTATCTGTGG 2139
D 7204 TCATTTAGCGAGATCGGCGAGAGACCACTTACTTCTGGGTGAGTTCACTTTGCAAG 7263
Q 2140 ACCGATTAACAGAGACGCGCTGAGATCGGAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
D 7264 ATTGCTTACCTCCACAGCGGTGCGGAGACCTTGAAGAGCTGTTTAAAGTTGG 7323
Q 2200 GAAAACCATTTGCGAGTGAATACCAAGACTGCGACCGCGCGGCACTGATGATG 2259
D 7324 GTAAACCGCTCCGACCGAGATGAGCAAGACGAAGACGAAGACGCGCTCTGCTAATG 7383
Q 2260 AAGCAATGCAATGGAACAAATTGAAATTACGACGAGTTAGTAAGCGCTGATACATCA 2319
D 7384 AAACAAAGGCGTGGTTAGAGTATTAACAGACACTTATGAGTGAGCGCTGGCAATC 7443
Q 2320 GATACGATCATCTGCGACGCGCTGATCATCAGTCTGTGCTCAAGCTTACCGGAAAGCG 2379
D 7444 GGTATGAGTATGACAACTACCTGCTGCTGAGATGAACTTTTGGCCAGAGCA 7503
Q 2380 TTAAAGACTTCANAGAGATTAAGAGAGCCCAATACCTTCAAGCTGATCTTAAATAG 2439
D 7504 AAGAGACTTTCAAGCATCAGAGGGGAAATTAAGCATCTCTACGTTGCTTAAATAGT 7563
Q 2440 TGACGTATGTA-----CAGCACTTACCACCGGAGAAATTTTCA 2482
D 7564 CAGCATGATCATTTCACTGACTAATATACCAACACACACCATGATGATGAGATTTCT 7623
Q 2483 TACCTCAGCTGAATTTCCACAGTTTAACTTACCAATCGGATGGCTTACCGAGATCCA 2542
D 7624 TTAAATGCTGCGCGCGCGCTTCCAGCCCACTGCTGATGAGCGCGGAGAA 7683
Q 2543 AACCTCTTAGGCGCGCTGAGAGCGTTTGGCGCCCGCTGGTGTCAATGAAGAT 2602
D 7684 GAGAGGAGGGGCGCGCATGCTGCCGCAATGGGCTGGCTTCCCAATCCAGCACTGA 7743
Q 2603 CTTAGAGGTTGATGATCACTTCAATCAAC--GATCACTTAATCGCGCCAG 2660
D 7744 CCACAGCGGTACGTGCTTGTGATGAGCAAGCACTAGACTCAAAACCCACGCGCAC 7803
Q 2661 GTCCACCGCCAAAGAAAGAGAGTGTCTTAGCCAAAACCTACAGCTTAAAGAA 2720
D 7804 GCCCGCGCGCGCGCAAGAGAGGCGCCAAAGCAACCAAGAGCGGAAGAACCA 7863
Q 2721 AGAAGCAGCAAGCCAGAGAGCAAAAC---GCAAGCTTAAACAGAGAAAGCAACGTA 2777
D 7864 AAACACAGGAGAAAGAAAGAGCACTTGCAAAACCAAAACCGGAAAGAGACGCTA 7923
Q 2778 TGTGTATGAAGTTGAGTGGAGCAAGCAATTTCCGATC--ATGCTGAACGGCAAGTGA 2834
D 7924 TGGCACTTAAGTTGAGGCGCAGACAGCTTTGAGCTCAAAAATGAGGACGGAATGCA 7983
Q 2835 ATGATATGCTGTGCTGCGAGAGAGGCTGATGAACCACTCAAGTTGAAGAAAA 2894
D 7984 TCGGGACGCACTGGCCATGAGAAAGAAAGTAATGAACCACTCCACGTGAAGAACTA 8043
Q 2895 TTGATTAAGCAATTAGCGCGCGCTGAATTTGAAGAGGCTAGACATGATGAGT 2954
D 8044 TTGACACACCTGTGCTATCAAGCTCAAAATTCACCAAGTCTCAGCATACGATGAGT 8103
Q 2955 ACGGCGAGTTTCCCAAGAACTGAATATGACAGCTGAGTACACGAGCAAACTCAC 3014
D 8104 TCGCACAGTTGCGGCTCAACATGAGAAAGTGAAGCGTTACCTACCAACAGTGAACCTTG 8163

Q 3015 CGGGCTTTTCAAACTGGCACACGCGCGAGTCCAGTATGAAATGGGAGATTTACCGTAC 3074
D 8164 AAGGTTTCAAACTGGCACACGAGAGCGGTGAGTATGATGAGGAGATTTACATCC 8223
Q 3075 CGAAGAGTGGCGGGGAAAGGCGACACGCGAAAGACCGATCTTGACAAAGAGCGAG 3134
D 8224 CCGCGGATGAGGAGGAGAGAGACAGTGTGCTCCGATTATGATTAATCTCAGGCGGG 8283
Q 3135 TTGTGGTATTTGTTTATGAGAGTGCAAATGAGGGCACGCTAAGCGGCTTTCAAGTGTCA 3194
D 8284 TTGTGGATATGCTCTCGAGGGGCTGATAGGAAACAAACCGCTTTTGGCTGTCA 8343
Q 3195 CTGGAACAGAAAGGGGATACCATTTAGGATACCCCGAAGTTTGAACCGGT--- 3251
D 8344 CTGGAATAGCAAAAGGAGAAACATCAAGCAACCCCGAAGGAGCAAGAAAGTGTCTG 8403
Q 3252 -----CACTAGTTACAGCGCTATGCGTCTTGAATGTCAAGTTTCCATGGCAAAAC 3305
D 8404 CTGACACACTGATCAAGCGCATGTGCTTGGAAACGTGAGCTTCCATGCAATCGCC 8463
Q 3306 CACCGTGTGCTATTTCACTAAGCCGAGAACGAACATCTGACGTGCTGAAGAAAGCTCG 3365
D 8464 CGCCCATGCTACACCGCGAAACATCCAGAGCTTCGACATCTCGAAGAGAAACGTGA 8523
Q 3366 ACAATCCAAATTACAGACGCTGAGAAACGTTGAAT---GTCCATACAGCGCGGC 3422
D 8524 ACCACAGGCGCTACAGACCTTGTCTCAAGCCATATTTGGGTGCGATCTCGGCGAGAA 8583
Q 3423 CCAAAAGAACATTTACCGATGACTTCACTGACCAAGTCTCTTACCTGGGCTTGCCTGT 3482
D 8584 GTAAAGAACGTCATCGACGACTTATCTTGAACAGCCGCTACTTGGGACATGCTCGT 8643
Q 3483 ATTGCAACACTCAAGCGCGTGTTCAGCCCAATAAATTTGAGAACTGTGGGAGCAAT 3542
D 8644 ACTGTCAACATCTGAACCGTGTCTTGAAGCCGATTAAGATGAGAGCTGTGGATGAAG 8703
Q 3543 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3602
D 8704 CGGACGCAACACCATACGATACAGACTTCCGCGGCTTGGATGAGCAACAAAGCGGAG 8763
Q 3603 CTGGGATGTCAACCAATTTCCGTTTACATGTTCTTTCGACAGACCATGATCAAGGAG 3662
D 8764 CAGCAAGCTCAAAATTAAGTACCGGTACATGTCGCTGAGAGAGATCATATGTCAAAG 8823
Q 3663 ACAGTATGAGAAATATGCTATCAGCACTTGGACCTTCCGCTGTGCTTGGCCCAAG 3722
D 8824 GCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8883
Q 3723 GGTACTTCTGTTAGCTCAATGTCCTCAGGTGACAGTGAACGTCAGTATCAAGAGCG 3782
D 8884 GATCTTTCTTCTCGCGAAGTGTCTTCAAGGAGACGCTGATGATGATGATGATGATGAT 8943
Q 3783 GAGCATGAGAAATTCATGACCGGTGAGAAAAAGATCAGAGAGAAATTTGTGCTGAG 3842
D 8944 GCAATCAGCAAGCTATGCAATAGCGCGCAAGATTAACCAAAATTTGTGAGAGGG 9003
Q 3843 AGGAGTACTTGTTCACACCGCTCATGAGAAAGCTGTGAAGTGCACAGTTTATGATCACT 3902
D 9004 AAAAATATGACTTACCTCCGCTTACAGTAAAGAAATTCCTTGAACGTATGACACGTC 9063
Q 3903 TGAAGAGAGCTGCGCGGGGTACATGACATGACAGAGCGGCGCACAGCGGTATAGT 3962
D 9064 TGAAGAGAGAGCGCGGGCTATCTATCTATGACAGGCGGAGACGCTATATCAT 9123
Q 3963 CTAATCTGAGAGAGCTGACGCGGAATGTATTAACCACTTCTGCAAGAGCTCA 4022
D 9124 CTAATCTGAGAGATCATTCAGGAAAGTTTACGGAAGCCACCATCGGGAAGAACTTA 9183
Q 4023 CTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4082
D 9184 CTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9243

OY	4083	A	CGGTGACATTAAGCAAAACAGTGCATTGGCTCAAAAGAGGACCAACGAAATGGGCT	4142
Db	9244	C	GGGTGACCGCCATCAAGAGTGGCTGGCTTAAGAGGACCAACGAAAGGGGCT	9303
OY	4143	T	CAACTCGCCGGAATCTTATTAAGGACACAGACCACTCAGTCAAGGTAAATTGACATTC	4202
Db	9304	T	CAACTCGCCGGAATCTCGATCAGACACGGCCACACCGCCCAAGGAAATTGCAATTTGC	9363
OY	4203	C	ATTCCGGTTGACACCGACAGTCTGGCCGGTTTCGGTTAGCTCACACGCTTACAGTCA	4262
Db	9364	C	TTTCAAGCTGATCCCGAGTACCTGATCGTATGCTTCCGTTGCCACGGCCGAACGTAGTAC	9423
OY	4263	A	GTGTTAAAGGCTACACCTTCCACTGACTGCATCGACCAACATTGCTGCAACGA	4322
Db	9424	A	CGGTTTAAACATCAGCCTCCCAATTGACACAGACCATCTGACATTGCTCAACCA	9483
OY	4323	G	AAAAATTGGGGCTGGACGACGCAACGAGAAATGGAATTACAGGGCTTACATCCAGA	4382
Db	9484	G	GAGACTAGGGGCAAAACCCGGAACCAACCACTGAATGATCATGGAACACGGTTGAA	9543
OY	4383	A	TTTTTCTGTGGGGAGAAAGGGCTGAGTACGTATGGGGTAAACATGAACAGTCAAG	4442
Db	9544	A	CTTCAACCGTGAACGAAATGAGCTGGAATCATATGGGGCAATCAGCAACAGTAAAGG	9603
OY	4443	T	CTGGGGCCAGAGATCGGCAACGAGCCACCATGATGGCCGATGATCATATCC	4502
Db	9604	T	CTATGCCCAAGAGCTGACACGAGAGACCTCTCAGATGGCCACACCAAAATGTAAGC	9663
OY	4503	A	CTATTATCATCGGATCAGCTACATGCTACATGTCATTTGTGCTGTGTGCTCTTGCTA	4562
Db	9664	A	TTACTATCATCGCATCTCTGTGTACACATCTTTAGCCGTGCATCAGCTCTGTGGGA	9723
OY	4563	T	CTCTGTAGGACATGACATCAGCAGCTTGATGSCCAAGACAAGAAAGACCTGCTGA	4622
Db	9724	T	GTAGTATGGGGTAACTGTGTGACGATTATGTGCTGTAAGCCGGCCGTGAGTGGCTGA	9783
OY	4623	C	GGCATACGCGCTTGACCGAAACGCAACGTAACCAACAGCATTAAGCGTTTGTGTGCA	4682
Db	9784	C	GGCATATGACCTGGCCCAATGCCGTAATCCATCACTTCCCTGGCACTTTTGTGTGTG	9843
OY	4683	T	TTCGGCCAAACGAGCTGAAACATTTGGAGAACTTTGAACATCTGTGTTTAAACAC	4742
Db	9844	T	TTAGTCCGCTAAATGCTGAACATTACCGGACACATGATTACTTATGTGTGCAACGCC	9903
OY	4743	A	ACCGTTTCTGTGGACAGATTGTGCAATCTCTGGACGCCCTTGTATTTCTGTTCGCT	4802
Db	9904	A	AGCGTTTCTGTGGGTCAAGCTGTGTATACCTGTGGCCGCTGTGCTTCTAATGCCCT	9963
OY	4803	G	CTTTTACATGCTCATGCTTTTTTATTTATGTTGACAGCGTGTGCTCTGGGAAAGTAAAG	4862
Db	9964	G	TTTGTCTACATGCTCCTGTGCTTTTTTATGTTGGTTGCGGGCTTACCTGGCAAGTAAAG	1002
OY	4863	C	CTTTCGAACATGCAACCACTGTGCCAAATGTTCCGGGGATACCGTATTAAGCGTTGGTGC	4922
Db	10024	C	CTTACGAACATGCAACCACTGTGCCAAATGTTGCCAAATACCGTATTAAGCACTTGTG	1008
OY	4923	A	AACGCGAGTTTACGCGCACTTAACTGAGATCAACGCTGTCTCATCGGAATTAAACAC	4982
Db	10084	A	AAAGGCGAGGTACGCCCGCGCTCAATTTGGAGATTACGTATGTCTCTCGAAGGTTTGC	1014
OY	4983	C	TTTCACTTAACAGAGATGAGCTTGCAAAATTCACACAGTCAATTCCTTCAACACAG	5042
Db	10144	C	TTTCACTTAACAGAGATGATTAACCTCAAAATTCACACATGTGTGCTCCCTCCCTTAAG	1020
OY	5043	T	TAAATGTGTGGGGTCCCTCGAATGCAAGGATCTCAAAGCGGATTAACATGCGCG	5102
Db	10204	T	CAGATGTGTGGGCTCTTTGGAATGTACGCCCGCGCTCAGCAGACATTAATCTGCAAG	1026
OY	5103	T	TTTTTGGGGTGTGTACCTTTTCAATGTGGGAGGGCACAATGCTTCTGTACAGTGA	5162
Db	10264	T	CTTTTGGAGGGGTGTACCCCTTTCATGTGGGAGAGCAATGTTTTTGTGCAACAGTGA	1032
OY	5163	A	CAACAACATGATGAGGCGTATCGTGAATTCGCTCAGACTGCAATATGATCAACGAC	5222

Db	10324	ACACCAGATGATGAGGCGCTACGTCGAATTGTCAGTGAATGCGGACCTGACCGACGCC	10383
Qy	5223	TCGCACAAAAAGTTTCACACAGCTGCTCTGAAAGTCGCGCCGCTATAGTATACGCGACA	5282
Db	10384	AGCGGATTTAAAGTGCATACCTGCCGATGAATAGAGCTGCGATTAGTGTACGGAAACA	10444
Qy	5283	CCACCGCGCACCTTGATATACGTTTGTCAATGGCGTCACGCCAGTTTCTCACGGACCTGA	5342
Db	10444	CTACCAAGTTTCTTGATGTGTACGTGAACGGAGTCAACACAGGAACGTCTAAAGACCTGA	10503
Qy	5343	AGGTCAATGACAGGCGCATATACGCCGCTTTTCACCCCTTTGACCAATAAAGTGTATCA	5402
Db	10504	AAGTCATAGCTGACCAATTTACGACTTGTTCACACATTTCCATACAAAGTGTATATCA	10563
Qy	5403	GAAGGGGCGCTGTTTACATACAGACTTCCCTGAGATAGAGGTATGAAGACGAGACGT	5462
Db	10564	ATCGCGGCTGTGTATCACTATGACTTTCCGGAATACGGAGCATGAATACGAGACGT	10622
Qy	5463	TCGGCGATATTTCAAGCACTCTCGCTGTATGCTACAGACATAGTAGCCGCACTGACATAC	5522
Db	10624	TTGGAGACATTTCAAGCTACCTCTTGATGCAAGACCTCATCGCACGACAGACATTA	10683
Qy	5523	GGCTGCTGAAGCCTTCTGTCAAGAACATCCAAGTCCCCCTACACCCAGCAATATCAGGCT	5582
Db	10684	GGCTACTCAAGCCCTTCGCGCAAAACGTGATCTCCGTAACAACGCGCGCATCTGAGAT	10743
Qy	5583	ATGAATGTGTAAGAAACATCACTACAGACGACCCCTGCAGAAACAGACCATTTGATGTA	5642
Db	10744	TCGAGATGTGTAAGAAACATCAAGCCGCGCCACTGACGAAACCGCCCTTTTGGGTGCA	10803
Qy	5643	AAATTTGAAGTGAAGCCTCTGCGAGCGTCTTAACTGTGCTTACCGGCAATCCCTATCTCA	5702
Db	10804	AGATTTGACATCAATCCGCTTCGAGCGGTGACGTCTATACGGGAACATTTCCATTTCTA	10863
Qy	5703	TTGACATTCCTGATGACGCTTTTGTGATCATCAAGATCAACAAATTTTGAAGTTA	5762
Db	10864	TTGACATTCGCGAACGCTGCTTTATACGACATCAAGATCACCACTGCTCTCAACAGTCA	10923
Qy	5763	GCTGCACAGTAGAGACATGCATTTATTTCTGCAACCTTTGGTGTCTTAACTTACAGT	5822
Db	10924	AATGTGATGTCACTGATGATGACTTATTTACGCGGACTTTCCGAGGAGATGGCTACCTGCAAT	10983
Qy	5823	ACAAAGCTGACAGGAGGAGCATGTGTCAAGTTCACTCCACATCCACAGACAGCTGTTTTGA	5882
Db	10984	ATGTATCCGACCGGGAAGGACAAATGCCCTGTATCATTCGATTCGACACAGCAACCTTCC	11043
Qy	5883	AGGAAGGACACACATGTGACTGCTCGGTAGGACGATTAACACTACATTTTACACATCGA	5942
Db	11044	AAGAGTCGACAGTTCACTGATCTGAGAAAGGACCGGTGACAGTACACTTCAGCACCGGCA	11103
Qy	5943	GCCCAACAAGAAATTTTATAGTTTCCGTATGGCGAAAGTCCACATCCAAATGCGAAT	6002
Db	11104	GCCCAACAGGGAACCTTATGTATCCGTGTGTGAAGAAAGCAATCCAAATGCGAAT	11163
Qy	6003	GTAACCAACCGCGCGACCATATATTTGAGAAACACATAAAGTCGACCAAGATTTCCAG	6062
Db	11164	GCAACCAACCAAGCTGATCATATGCTGAGCACCCCGGCAAAATATGACAAAGATTTCCAG	11223
Qy	6063	CGGCAGTTTCAAAACATCTTGAACTGAGCTGCTTSCACTGTTTGGGGGAGCATCC	6122
Db	11224	CCGCACTCTCAAAAACCTTCATGAGATGGGCTGTTTCCCTTTTCGCGCGCGCTCGTCCG	11283
Qy	6123	TCAATTTGTATGACATTAATAGTGTGTCTGACGCTCTATGCTTATAAACACAGTATAT	6182
Db	11284	TATTAATTTATAGCACTTATATATTTTGTTCGACGATGATGCTGATACACAGAAAT	11343
Qy	6183	GACTGACCGGCGACCTGA-CATAGCGGTAAACCTGATGTACTTCCAGGAAGCGTGT	6241
Db	11344	GACCGCTTAACGCCCAATGACCCGACCAAGAAACTGATGTACTTCCAGGAACGTGAT	11403
Qy	6242	GCATTAATGCCACGCGCGCT-----TGAAGT 6268	

Db 11404 GCATATGATCAGCTGTATATTAGATCCCGCTTACCGGGGCAATATAGCAACACC 11463
Qy 6269 AAAAAGTATGATTTCCGAGAGCAACGTCATATGCTGTGCAGCT 6318
Db 11464 AAAAAGTATGATTTCCGAGAGCAACGTCATATGCTGTGCAGCT 11513

RESULT 8
AR042409
LOCUS AR042409 11663 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5811407.
ACCESSION AR042409
VERSION AR042409.1 GI:5962905
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11663)
AUTHORS Johnston,R.E., Davis,N.L. and Simpson,D.A.
TITLE System for the in vivo delivery and expression of heterologous
JOURNAL Patent: US 5811407-A 1 22-SEP-1998;
FEATURES
source Location/Qualifiers
1..11663
/organism="unknown"
BASE COUNT 3291 a 3105 c 2907 g 2360 t
ORIGIN

Query Match 36.9%; Score 2396; DB 6; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

Qy 580 ACTGACGGTATGAGCGGAGCGTATATTTCTCATGGGAAACGGCCAGTACCTTC 639
Db 5707 ACTGTCACCGGGATGAGTGGTACATATTTGACGAGACAGCCCTGGSCACTTC 5766
Qy 640 AACGAATCAGTACGTCAATGTAATCAACAGAACCTATATGATGGGCGCTGCATG 699
Db 5767 AAAAGAAAGTCCGTTCTGAGAACAGCTTACAGAACGACCTTGGAGGCAATGTTCTGG 5826
Qy 700 AGAAGTATACGCCCCCGCTCGATCTCGAAGAGAAATGTTACAGAAAGTCTGC 759
Db 5827 AAAAATATCAGCCCGCTGCTCGACACGTGAAAGAGAACACCTCAACTCAGTAC 5886
Qy 760 AATATGCGCTCTGAAGAAATAGAGAGATATCATACGAAAGTAAATATGA 819
Db 5887 AGATGATGCCACCGAAGCCAAAGACAGTACAGTCTCGAAAGTAAATACCAAGA 5946
Qy 820 AAGCAATTAACAGCGAGGAGTCAATTTCTGATTTGGGCAATATCTATCATCAGAAGTGA 879
Db 5947 AAGCATTAACCACTGAGGAGTCTTTTCAAGGCTTACAGCTGTATTAAGCTTGCACAG-- 6063
Qy 880 ATCTGTGAGTGTATACAGATCAATTAATCTGTACCAATTAATCTGTCAACGTAATTA 939
Db 6004 ATCAGCGAGATGCTATTAAGATCACTAACCCGAACCACTGATTTCCAGCAGTAC 6063
Qy 940 ACAGGTTTAACTGCGAGAGTGGCGGTTAAAGTGAATTAAGTATCAAGAGAT 999
Db 6064 GAAATCTCTGACCCAAAGTTTGTCTGAGTGTGTTGAACACTATCTGTGATAGAA 6123
Qy 1000 ACCCTAAGTACGACGATATATGTAATACAGATGAATAGATGCTATCTGATGATG 1059
Db 6124 ACCGAGCGTATGATCTTATACATACAGAGTACAGATGCTTACTTGATATGATG 6183
Qy 1060 ACGGCGATCTGTCTGTATGATACAGCACTTTTGTGCGGCTAACTGGAAGTACC 1119
Db 6184 ACGGAGAGTGTGCTGTATGATCTGCACTTTTGTGCGGCAAGTGTGAAGTACC 6243
Qy 1120 CAAGAAGCATAGTATTTGAGCAGAGATAGATACGCGTCCATCGCTATACAGA 1179
Db 6244 CGAAAGACAGAGTATAGAGCCCAACATCCGAGTGCCTTCATCAGCGATGAGA 6303
Qy 1180 ATACATTACAAATGTATTTGCTGACGCTATAAAGGAATGCAACGTTACCAATG 1239

Db 6304 ACAGTTGCAAAAGCTCATTTGCGCGACTAAAGAACTGCAACGTCACAAATGC 6363
Qy 1240 GAGATTAATCTGCTTCTGATTTCCGCGGCAATTTATGTTGTTTCAAGAAATACCAT 1299
Db 6364 GTGAATCGCAACACTGAGACTCAGCGCATTTCAACGTTGATGCTTTGAAATATCAT 6423
Qy 1300 GCATGATGAGTACTGGGATACCTTTGCGATTAACCTATTCGGCTACTACAGAGACG 1359
Db 6424 GCATGACGAGTATTTGGAGAGAGTTTCCCGAAAGCCAAATTAGATACATAGATTCG 6483
Qy 1360 TTACGCAATATGTGACAAAGCTGAAAGGCGGAAAGCAGAGACATGTTTGGAAATCTC 1419
Db 6484 TTACCGCATAGTGGCCAGACTGAAAGGCTTAAGGCGCGGCACTGTCGAAAGACGC 6543
Qy 1420 ATATATCAAAACGCTTGCAGAGATACCAATGATGATCAATTCGTATGATCTAAAGAG 1479
Db 6544 ATATATTTGTCCTCATTCGCAAGAAAGTCCCTATGATATTCGTATGACATGAAAGAG 6603
Qy 1480 ATGTCAAAGTTATCCCGGACGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
Db 6604 ACGTGAAGTTTACCTGCGACGAAACACACAGAAAGAAACCGAAAGTACAGTGTATC 6663
Qy 1540 AGGCTGCAATCCCTTGTCTACCGCTTACCTTTGCGGATTCATCGGAAATTAAGTCCGTA 1599
Db 6664 AAGCGCAGAAACCCCTGCGACGCTTACCTATGCGGATTCACCGGAGTTAGTGC 6723
Qy 1600 GACTGAATGGGCTCTTCTGCAAAATTCATATCTCTTTCGACATTCGACGGAATAT 1659
Db 6724 GCGTTAAGCGGTTTGTCTACCAAACTTACACGCTTTTGAATGTCGCGGAGACT 6783
Qy 1660 TTGATGCGATTTATTTGCAACATTTCCACAGCGGACCGCAAGTATGGAACGACATCG 1719
Db 6784 TTGATGCAATCATAGCAAGAACCTTCAAGCAAGTGAACCGGATCTGAGACGATATCG 6843
Qy 1720 CGTGTGTTGTAAGAAAGCAAGACGAGCTATTCGCAATTTGCGGCTGATGATCTTGAAG 1779
Db 6844 CTTGTTGCAAAAGCCAGAAAGCAGACGCTATGCGGTTAAACCGGCTGATATCTTGAAG 6903
Qy 1780 ACTTATGCTGACCAACCCCTCTTATGATTTAGAGAGGCGGCTTGGCAATTCAT 1839
Db 6904 ACTGAGTGTGAGCAACCACTACTGACTTGTATGAGTGCCTTTTGGAAATATCAT 6963
Qy 1840 CTGTGACCTACCTACGAAAGAGGTTTAAATTTGGTGCATGATGAATCCGATATGT 1899
Db 6964 CCACCATCTGCGCACGAGTACCGCTTTCAATTTCCGGGAGATGATGAATCCGATATGT 7023
Qy 1900 TCTTAAAGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGATTAACG 1959
Db 7024 TCTCAGGCTCTTTGTCAACACAGTTCGATGTCGTTATCGCAGAGATATTTGAGG 7083
Qy 1960 AACGTTAACCAAGTACGCTGCGCGGCTCTATTCGCGAGCAAGTAATAGTATGATG 2019
Db 7084 AGCGGCTTAAAGCTTCAAAATGTGACATTTATTCGCGAGCAACAAATTAATACGAG 7143
Qy 2020 TCGTCTCCGACACTTATGAGCGGAGAGATCGGCACTTGGCTGAACATGAAAGTAA 2079
Db 7144 TAGATATCGAAGAAATGAGCTGAGAGTGTGCACTGGCTCAATGAGATGATTAAGA 7263
Qy 2140 ACCAGATTAACGAGCAGCGCTGAGAGTCCGCAACCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7264 ATTCGTTTACCTCCACAGCGTGTGCTGCGGAGCCCTTGAAGAGGCTGTTTAAAGTTGG 7323
Qy 2200 GAAACCATTTGCGAGTATGATACCAAGACTGCGACCGCGCGGCACTGCATGATG 2259
Db 7324 GTAAACCGCTTCCAGCGGACGATGAGCAACAGACAAAGACGCGCTCTGCTAATG 7383
Qy 2260 AAGCAATGCGATGGAACGAATTTGAATTAACGAGGTTAGTAAAGGCTGTAATCA 2319

Dh 7364 AAACAAGCGCTGTTAGAGTAGGTATPACAGACACTTATAGCACTGCGCGTGGCACTC 7443
Qy 2320 GATACGATATATATGCGAGCGCTGATCATCATCTCTCTGCACTGATGCGGAAGG 2379
Db 7444 GGTATGAGTAGACAACTACACCTGTCTGTCTGGCATTTAGGAATTTTGGCCAGAGCA 7503
Qy 2380 TTAAAGACTTCAAGAGCATPAGAGGAGCCCAATCACTCTTAAGGCTTGAAGCTTAATAG 2439
Db 7504 AAAGAGCAATTCAGACCATCAGAGGGGAAATTAACATCTCTACGCTGTCTTAATAGT 7563
Qy 2440 TGAAGTATAG-----CAAGCACTTACCCACCGGAGAAATGTTTCA 2482
Db 7564 CAGCATAGTACATTTCACTGATCTAATACCAACACCACTCAATGATAGGATCT 7623
Qy 2483 TACCTCAGCTGAATTTTCAACCACTTACCTTAATTCGATGCTTACCGAGATCCA 2542
Db 7624 TTAACATCTGCGCGCGCGCTTCCAGCCCCCACTGCACTGAGAGCGCGGAGAA 7683
Qy 2543 AACCTCTAGGCGCGCTGAGGCGTTTCGGCCCCGCTGGCTGCTCAATCGAAGAT 2602
Db 7684 GAGGCGAGCGCGCGCATGCTGCGCCGCAATGGGCTGCTTCCCAATCCAGCACTGA 7743
Qy 2603 CTTAGAGAGTGAATGATCACTTGACTTTCAAAACAC--GATCACTTAATCGCGCCAG 2660
Db 7744 CCACAGCGGTGACCTTACCTTATGATGACAGCACTAGACCTTAACCCCACTGAC 7803
Qy 2661 GTCCACCGCAAGAAAGAAAGAGTGTCTTAACTCAAAACCTTACGCTTAATAAGA 2720
Db 7804 GCCCGCGCGCGCGCGCAAGAAAGAGCGGCCCAAGCAACACCGAAGCGCAAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCAAGAGAGCAAAAC--GCAGCTTAAACAGAGAAACGACACCTA 2777
Db 7864 AAACACAGAGAAAGAAAGAAAGCACTGCAAAACCCAAACCCGAAAGAGACGCTA 7923
Qy 2778 TGTGATGAATTTGAGTTCGACAAAGATTTCCGATC---ATGCTGAACCGCCAGTGA 2834
Db 7924 TGGCACTTAAGTTGAGGCGCGCAACCTGTTGACAGTCAAAATAGAGACGAGATGCA 7983
Qy 2835 ATGATATGCTGCTGTGCGAGAGAGGCTGATGAACCACTCCAGCTTGAAGGAAAA 2894
Db 7984 TCGGCGACGCACTGCGCATGGAAGAAAGATTAATTAACCACTCCAGCTGAAGGAACTA 8043
Qy 2895 TTGATATGAGCAATTAAGCGCGCGTGAATTTGAAGAGCTAGCATGTAACGATTTGAGT 2954
Db 8044 TTGACCACTGTGCTATCAAGCTCAATTCACCAAGTCTGACATTAACGACATGAGT 8103
Qy 2955 ACGGCGAGTTCCTCCAGAACTAGAAATCAGACCGCTGCACTACCGACGAAACCAAC 3014
Db 8104 TCGACAGATTCCTGCTCAAGATGAGAGGCTTCACTACCACTGCAAGCAACCTG 8163
Qy 3015 CGGCTCTTACAACTGGACCAAGCGCGAGTCCAGTATGAGAAATGGAGATTTACCGTAC 3074
Db 8164 AAGGTTTAACTGGAACAGAGGCGGTGCAATATGAGAGGCAAGATTTACATCC 8223
Qy 3075 CGAAGAGTGGCGGGAAGAGCGCAAGCGGAAGACGATCTTGAACCAAGAGGAGAG 3134
Db 8224 CCGCGGAGTGAAGAGGAGAGAGCAAGTGTGCTGCAATATGATTAACCTAGCGCGG 8283
Qy 3135 TTGAGCTAATTTGTTAGAGAGTGAATGAGGCGACCGCTTCAAGTGTCA 3194
Db 8284 TTGTGCGCATATGCTCTCGAGAGGCTGATGAGGGAACCAAGAACCGCTTTCGCTGCA 8343
Qy 3195 CTTGGAACCAAGAGGAGTGAACATTAAGGATACCCCGGAAGGTTCTGAACCGTGT--- 3251
Db 8344 CTTGGAATAGCAAGAGGAGAAACATCAAGACACCCCGAAGGAGCAAGAGTGTGTG 8403
Qy 3252 -----CACTAGTTACAGCGCTATGCTGCTTCAATGTCAAGTTCCTCATGCAAAAC 3305
Db 8404 CTGACACACTGTGTACCGGCAATGTGCTTGTGAAGAGTGAAGTTCCTCATGCAATGCC 8463
Qy 3306 CACCGGTGTCTTATCTAGAGCGCAAGACGAACTGACGCTGCTGAAGAGAACTGTG 3365
Db 8464 CGCCCACTGTCTACCGCGCAACCATCAAGAGCTCTGACATCTCTGAAGAGAACTGTA 8523

Qy 3366 ACAATCCAAATTAAGACACGCTGTGAGAAAGTCTTGAAT---GTCCATACCGCGGC 3422
Db 8524 AOCACGAGCGCTACGACACCTGTGCTCAACGCGCATATTCGGGTGCGGATGTGCGGAGAA 8583
Qy 3423 CCAACGAGCATTAACGATATCTTACATGACCTGACCTTACCTGCGGCTTTCGCGCT 3482
Db 8584 GTAAAGAGGCTCACTACGACTTATACCTTGAACGAGCCGCTTACTTGGGCACTGCTGT 8643
Qy 3483 ATTGCAGACTCAACGCGGTGTTACGCCCAATTAATTAAGAAAGCTGTGGGCGAAT 3542
Db 8644 ACTGCACCAATCTGAACCGTGTCTTACCGCTGATTAATGAGAGGCTGTGGGATGAAG 8703
Qy 3543 CTGATGATGATGATTAAGATTAAGAGTCTGCGCAATTCGCGCTTACATGACGAGCA 3602
Db 8704 CGAGAGCAACCATACGATACGACTTCCGCGCACTTGTGATACCAACCAAGCGGAG 8763
Qy 3603 CTGGGATGTACCAATTCGCTTACATGCTTTTGACCAAGACCATGACATCAAGAG 3662
Db 8764 CAGCAAGCTCAAAATTAAGTACGCTACATGTGCTGACAGAGATCATCTGTCAAAAGAG 8823
Qy 3663 ACAGTTAGGAGAAATTAAGTATACGACATCTGACCTGCGCTGTCTTGGCCCAAG 3722
Db 8824 GCACATGATGATGACATCAAGATCAGACCTGAGACCGTGTAGAGGCTTATGCTTAAAG 8883
Qy 3723 GGTACTTCTGTAGCTCAATGTCTCTCAGGTTGACAGTGAACCGTCACTATCAAGAGC 3782
Db 8884 GATACTTCTCTCGGCAAGTGTCTCTCAGGAGACAGGTTAAGCTTATGATAGGATGA 8943
Qy 3783 GAGCATCTGAATTCATGACCGGTGAGAAAGATCAGAGAGAAAGTTGTGCTAGAG 3842
Db 8944 GCAACTGACCAAGCTCATGACCAATGCGCCCAAGATTAACCAAAATTCGTGGACGG 9003
Qy 3843 AGAGTACTTGTTCCTCCAGCTCCATGGAAGCTGTAAATGCTTACGATCACT 3902
Db 9004 AAAATATGACCTTCTCCGTTACAGGTGAAGATTCCTTGAACATGTACGACGCTC 9063
Qy 3903 TGAAGAGAGCTGCGCGGTACATTAACCATGACAGGCGCGGACCAAGCGTAAAGT 3962
Db 9064 TGAAGAGAAACACCGCGGCTACATCACTATGACAGGCGCGGACCGATGCTTATCAT 9123
Qy 3963 CTTATCTGAGAGAAAGCTCAGGCGAAGTGTACATTAACCACTTTCGGAAGACGTC 4022
Db 9124 CTTATCTGAGAGAAATCATCAGGAAAGTTTACGGAAGACCACTCCGGAAGAACTATTA 9183
Qy 4023 CTTACGATTTAATGTGTGCGCACTACAGACAGGTATGTGAGCAACGGAAGATGA 4082
Db 9184 CTTACGAGTGTCAAGTGTGCGGCGATTACAAAGACCGGAACCGTTACGACCCGTAACG 9243
Qy 4083 ACGGCTGCACTAAAGCAAAACAGTGCATTGCTTCAAGAGGAGCAAAAGGAGTCT 4142
Db 9244 CGGCTGTACCGCATCAAGAGAGTGTGCTTCAATTAAGGAGCAAAAGAGTGTGCT 9303
Qy 4143 TCAACTGCGCGGATCTTATTAAGGACACAGACCACTCAGTCAAGTAAATTTGACATTC 4202
Db 9304 TCAACTGCGCGGATCTTATTAAGGACACAGACCACTCAGTCAAGTAAATTTGACATTC 9363
Qy 4203 CATTCCGTTTACACCGCACTGTGCGCGGTTCCGTTAGCTTACACGCTTACAGTACGA 4262
Db 9364 CTTTCAAGCTGATTCGAGTACCTGCACTGTGCTTGTGCTTACAGCGCGCGCAAGTATAC 9423
Qy 4263 AGTGTTCAAAGGATATCACTTCACTGACCTGCACTGCAATGCAAACTGTGCAACAG 4322
Db 9424 ACGGCTTAAACATCACTGCTTCAATTAAGACAGACCATCTGACATTTGCTACACCA 9483
Qy 4323 GAAATTTGGGCTCGAGACAGCAACAGAGAGATGATTAAGAGGCTCTACATCTCAGGA 4382
Db 9484 GGAAGCTAGGGGCAAAACCGGAACCAACCTGAATGATATGGAACACGCTTGAAG 9543
Qy 4383 ATTTTCTGTGGGCGGAGAGGCTGAGATGATGATGGGTTACATGAACCACTGAGAG 4442
Db 9544 ACTTCAACGCTGACCGAGATGCTGGAATTAATATGAGGCGCAATCAAGAACCACTGAAG 9603

OY	4443	TCGTGGCCCCAAGAGAGTCGGCACACAGGCGACCCCAATNGATATGGCCGATAGATATCATATCC	4502
Db	9604	TCCTATGCCCCAAGAGCTCTGACCACGAGAGCCCTCAGATGGCCACACCAATTAAGTACAGC	9663
OY	4503	ACTATTAATCATTCGGGATCCAGTCTTACATGTGTCATGTGTCGTGTGTCCTCTTGCTTA	4562
Db	9664	ATTACTATCATTCGGCATCTCTGTGTATACATCTTTAGCCGTGGCATATCACTGTCTGTGGCGA	9723
OY	4563	TCCTGTGTAGGCATCTGCATCATTCAGCAGCTTGCGATGCCAAGAGCAAGAAAGAGACTGCTGA	4622
Db	9724	TGATGATTTGGCGTAACTGTTCAGCATTAATGTGCTGTAAAGCGCGCGTGAAGTCCCTGA	9783
OY	4623	CGGCATATAGGGCTTGACCGGAAGCAAGGATACCCAGCATTTAGCGGTTTTGTGTGTCGA	4682
Db	9784	CGCCATATAGCCCTGGCCCCCAATGCGGTATTCCACTTCCTGGCACTTTTGTGTGTGTG	9843
OY	4683	TTGGGCAACAACGCTGAAACATTTTGGAGAACTTTGAACCATCTGTGTGTTTAAACAAC	4742
Db	9844	TTAGGTGGGCTTAATCTTAATACATTACCGGAGACATATAGTTACTTAATGTGTGAACAGCC	9903
OY	4743	AACGGTTTCTGTGGGCAGATTGTGCATCTCTGTGGACGCGCTTGTTATTTCTGTTCCGCT	4802
Db	9904	AGCGGTTCTTCTGGGTCACGCTGTGTATATACCTCTGGCCGCTGTGTGCTTTAATAGCGCT	9963
OY	4803	GCTTTTCACTGCTGCATGCTTTTATTTATGTGTGCAGGCGCTCTGCTGGGGAAAGTATACG	4862
Db	9964	GTTGTCTATGTGTGCTGCTGCTTTTATTTAGTGTGGTGGCGGCGCTTACTGTGGCAAGGTATGACG	1002
OY	4863	CCTTGCAACATATGCGACCACTGTGCGCAATATGTCCGGGGATCCCGTAATAGGCGTTGGTGC	4922
Db	10024	CCTATGAAACATATGCAACCACTGTTCCMAATGTGCCACAGATATCCGTATTAAGGCACTTGTTG	1008
OY	4923	AACGGCAGTTTACGCGCCACTTAACTGTGAGATCACGSGTGTCTCATTCGGAATTAACAC	4982
Db	10084	AAAGGGCAGGGGATACGCCCGCTCAATTTGGAGATTACTGTATGTCCTCGAGGTTTTGC	1014
OY	4983	CTTCACTATCAACAAGAGATACGTGACCTTGCAATTTCCACACAGTACTTCTTACCACACAG	5042
Db	10144	CTTCCACCAACCAAGAGATCAATTACCTGTGCAAAATTCACCACTGTGTGTCCTCCCTTAAG	1020
OY	5043	TTAATATGTGCGGGGTCCTCGAGTGCAGAGGCAATCCTCAAAAGCGGATTTACCATATGCCCGC	5102
Db	10204	TCAGATGCTGGGGCTTCTTGTGAAATGTACGCCCGCTCACGACGACATATCTGTGCAAGG	1026
OY	5103	TTTTTGGCGGTGTATCCCTTTTCATGTGGGAGGCGCAATGCTTCTGTGACAGTGAAGA	5162
Db	10264	TCTTTGGAGGGGTGTATCCCTTTCAATGTGGGGAGGACAAATGTTTTTGGCAGAGTGA	1032
OY	5163	ACAACAACAATGATGAGGCGGTATCGTTCGAGTTCCGTCCAGACTGCATATATATACCGAC	5222
Db	10324	ACAGCGCAATATAGTGAAGGCGTATCGTCAATTTGTCAATGTATGTGGCGCATCACACGCGC	1038
OY	5223	TGCGACTAAAGTTTACACAGCTGTCTGAAAGTGGGCTGCGTACGTATATATAGGCAACA	5282
Db	10384	AGGCGATTTAAGGTGCATCTGCGCGATGAATGAGACTGCGTATATGTATGATCGGGATCA	1044
OY	5283	CCACGCGCACTGGATACGTTTGTGCAATGGCGTTCACGCCAGTGTCTTCAACGGGACCTGA	5342
Db	10444	CTACCAAGTTTCTATGATGTGTATCGTGAACGGAGTACACACAGGAACGTCTAAAGACTGA	1050
OY	5343	AGGTCTATAGCAGGGCCGATATATAGCCGCTTTTCACTCTTTGACCATTAAGTGTCTATCA	5402
Db	10504	AAGTCATATAGCTGAGCAATTTTCAGCATTTGTTTACATTCGATATCAACAAGTCTGTTATCA	1056
OY	5403	GAAAGGGGCTTTTATCAACTATACACTTCCCTGAGTATGAGGCTATATGAATAACAGAGGCT	5462
Db	10564	ATCGCGGCTGTGTACAACTATATGACTTTCCGGAATATCGAGGCAATGAACAGAGGCT	1062
OY	5463	TCGGGATATTTCAAGCATCTCTGCTGTATGCTACAGACATATGATAGCCCGCATGACATAC	5522
Db	10624	TTGGAGATTTCAAGCTATCTTCTTGAATTAAGCAAGACCTCATGCTCCAGCACACAGCATTA	1068
OY	5523	GCGTGTGAGACCTTCTGTCAAGAACATTCAGTCCCTCTACACCCCAAGACATATACGGT	5582

Db	10684	GGCTAATCTCAAGCTTCCGCCAAGACGTGATGTCCCTTACACCGACGGCCGCACTCTGGAT	1074		
QY	5583	ATGAAATGTGGAAACAACAACTCAGAGCGACCCCTGCAGAAAGACAGCCATTGGATGTGA	5642		
Db	10744	TCGAGATGTGGAAAAACAATCAGACGCCGCCACCTGCAGAGAAACCGCCCTTTTGGGTGCA	1080		
QY	5643	AAATTGGAAGTGAAGCTCTCTGGAGACCTCTTAACTGTGTCTTACGGGCACATCCATCTCGA	5702		
Db	10804	AGATTGGAAGTCATACCTCCCTCTGACCGGTGACCTGCTATACGGGAACATTTCCACTTCTTA	1086		
QY	5703	TTGACATCCCTGATGACAGCTTTTGTGATCATCAGATACCAACAATTTTAGAAATTA	5762		
Db	10864	TTGACATCCCGAAGCGTCCCTTTATCAGACATCAGATGACCACTGCTCTCAACAGTCA	1092		
QY	5763	GCTGCACAGTAGACAGACTGCAATTTATTTCTGACACTTTGTGTGTTCTTAACATTACAGT	5822		
Db	10924	AATGTGATGTGATGAGTAGGACCTTATTACAGGGACCTTCGAGGGATGGCTAACCTGAGT	1098		
QY	5823	ACAAAGCTGACAGGAGGAGGACATTGTCTCACTTCTCCACTCCACAGACAGCTTTTGA	5882		
Db	10984	ATGTATCCGACCCGGAAAGACAATGCCCTGTATCTTGACATTTGAGACGACAACCTCC	1104		
QY	5883	AGGAAGGACACACATGATGACTGCCGTAGGACGATPAACACTACATTATTTAGACATCGA	5942		
Db	11044	AAGAGTGACAGTTTCAATGTCTCTGGAGAAAGAGCGGTGACAGTACCTTACAGACCGCGA	1110		
QY	5943	GCCCAACAAGCAAAATTTTATAGTTTCTGTATGCGGCAAGAAAGTCCACCTGCATGCTGAAT	6002		
Db	11104	GCCCAAGGGGAACTTCATTGTATTCGTGTGTGTGTGAAGAAACAATGCAATGCAGAT	1116		
QY	6003	GTAAACCAACGCGCGGACCAATTAATTGGAGAACACATAAAGTGCACCAAGATTTCCAG	6062		
Db	11164	GCAAAACCAACGCTGATCATATCTGTGACACCCCGCAAAAAATGACCAAGATTTCCAG	1122		
QY	6063	CGGAGTTTCCAAACATCTTGGAACTGGGCTGCTGTGACTGTTTGGGGAGCATCATCC	6122		
Db	11224	CGGCATCTCAAAACTTCAATGAGCTTGCTGTGTTGCCCTTTTCGGGGCGGCTCTCGTCC	1128		
QY	6123	TCATTGTGTGAGACTTATAGTGTGTGTCTGCAGCTTATGCTTATTAACACAGTAGAT	6182		
Db	11284	TATTAATTAATGAGCTTATGATTTTCTCTTGACAGCATGATGCTGACTAGACACAGGAAT	1134		
QY	6183	GACTGAGCGGCGGACACTGA-CATAGCGGTAAAACTCATGTACTTCCGAGAAAGCGTGT	6241		
Db	11344	GACCGCTACGCCCCAATGAGCCCGACGACAAAACTCGATTACTTCCGAGGAATGATGT	1140		
QY	6242	GCATPAATGCCACGGCGCGCT-----TGACACT	6268		
Db	11404	GCATPAATGCATCAGCGCTGTATATTAGATCCCGCTTACCGCGGGCAATATGACACACC	1146		
QY	6269	AAAATCGATGTATTTCCGAGAAAGCAAGTGCATTAATGCTGTGCACTGT	6318		
Db	11464	AAAATCGACTATTTCCGAGAAAGCCAGTGCATATGCTGGCGCACTGT	11513		
RESULT 9	AR096555	11663 bp	DNA	linear	PAT 08-SEP-2000
LOCUS	AR096555				
DEFINITION	Sequence 1 from patent US 6008035.				
ACCESSION	AR096555				
VERSION	AR096555.1	GI:10025463			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 11663)				
TITLE	Johnson, R.E., Davis, N.L. and Simpson, D.A.				
FEATURES	System for the in vivo delivery and expression of heterologous genes in the bone marrow				
FEATURES	Patent: US 6008035-A, 1 28-DEC-1999;				
SOURCE	Location/Qualifiers				
	1..11663				

BASE COUNT 3291 a /organism="unknown"
ORIGIN 3105 c 2907 g 2360 t

Query Match 36.9%; Score 2396; DB 6; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

QY 580 ACTGACGATGAAACGGGAGCTATATTTCTCATCGGAAACAGGCCAAGCTCACTTC 639
DB 5707 ACTGCTAACCGGGGTAGGTGGTACATATTTTTCAGCGACAGGCCCTGGGCACTTGC 5766
QY 640 AACGAAATCACTACGTCAATGTAACTACAGAACTATATTTGATGGGCGCTGCATG 699
DB 5767 AAAAGAACTCGCTTCTGCAAGAACGACTTACAGAACGACCTTGAAGGCAATGTTCTGG 5826
QY 700 AGAAGTATTAGCCCGCGCTCGATCTCGAAAGAGAAATGTTTACAGAAATCTGC 759
DB 5827 AAAAGATCTAGCCCGCGTGTCTGACAGCTCGAAAGAGAAACAGCTCAAACTCAGGTAC 5886
QY 760 AATTATGCGCTCTGAGAAATAGAGAGGTATCAATCCAGAAAGTAGAAATATGA 819
DB 5887 AGATGATCCCAACGAGCCCAACAAAGAGGTACAGTCTGAAAGTAGAAACCAAGA 5946
QY 820 AAGCAATTACAGCGAGGAGCTCATTTCTGATTTGGGACATATCTATCATCAGAACTGA 879
DB 5947 AAGCATTAACCACTAGGAGCTGCTTTCAGGGCTACGACTGTATTAACCTGTCCACAG--- 6003
QY 880 ATCTGTGAGTGTATACAGAGTCAATTAATCTGTACCAATCTACTGTCAACGGTAATTA 939
DB 6004 ATCAAGCCAGAAATGCTATTAAGATCACCTAACCCGAAACCAATCTATTTCCAGAGTATCAC 6063
QY 940 ACAGGTTACATCTGCAAGAGTGGGGTTAAACGTGCACTTAATTTCAGAAAGATT 999
DB 6064 CGAATCTCTGAGCCCAAGTTTGTGTAGCTGTTTGTAAACAATCTGATGAGAAATT 6123
QY 1000 ACCCTACAGTACCGAGTATTTGTATACAGATGAATGATAGATCGTATCTTGACATGTGG 1059
DB 6124 ACCGACGGTATGACATTTATGATCAGATCAGCAGATGATGATCTTATGGAATATGTAG 6183
QY 1060 ACGGCGCATCGTCTGTCTAGATACAGCCACTTTTGTCCGGTAAACTGAGAACTTACC 1119
DB 6184 ACGGACAGTGGCTGTCTAGATACAGCACTTTTGTCCCGCAAGCTTAGAAATGTACC 6243
QY 1120 CAAGAAGACATAGCTATTTTGCAGCCAGATGAAGTACAGCGTCCCATCTGCTATACGA 1179
DB 6244 CGAAAGACACAGAGATAGAGCCCAACATCCGAGTGGCTTCCATCAGAGATGACGA 6303
QY 1180 ATACATTCAGAAATGTATTTGGCTGAGCTATTAAGAAATGCAACGTTACCCAAATGC 1239
DB 6304 ACAGCTTCCAAACGTGCTCATTTGCCGAGCTAAAGAAATGCGACGTACACAAATGC 6363
QY 1240 GAGAAATTACCTGTCTTGAATTCGGGCGCATTTAATGTTGATTGTTCAAGAAATACGAT 1299
DB 6364 GTGAATCTCCAAACATGACTCAGGACATTCAGATGATGATGATGTTGAAAATATGAT 6423
QY 1300 GCAATGATGATGATCTGAGATACCTTTCCGAGATTAACCTATTCGCTTAATCAGAGAACG 1359
DB 6424 GCATAGAGAGATTTGGAGAGATTTGCCGAAACCAATTAAGATCTACTATGTAGTTGG 6483
QY 1360 TTACGCAATATGTGCAAAAGCTGAAGAGGCGGAAAGCAGACATTTGTTGGGAATATCTC 1419
DB 6484 TTACGCGCTAGCGTGGCCAGATCTGAAGGCCCTTAAGGCCGCGCATGTTCCGAAAGAGC 6543
QY 1420 ATTAATCTAAAACCGTTGAGAGATACCAATGATCAATTCGTCATGATCTTAAAGAGAG 1479
DB 6544 ATTAATTTGCTCCATTTGCAAGAAAGTCCATATGATAGATTGTCATGAGCATGAAGAGG 6603
QY 1480 ATGTCAAGTATCTCCGCGACGAAACATACAGAGAGCGCTTAAAGTGCAGGTTATTC 1539
DB 6604 ACGTAAAGATTAACACTTGCGAAACACACAGAAAGAAAGCCGAAAGTACAGTGAATAC 6663
QY 1540 AGGTCGAGATCCCTTGCTACCGCTTACCTTTGGGGATTCATCGGGAATTAATGTCGTA 1599

DB 6664 AAGCCGAGAAACCCCTGGCGACCGGCTTACTATGCGGATCCACCGGGAGTTAGTCCGA 6723
QY 1600 GACTGAATGCGGTGCTTTGCGCAATATCAATCTCTTTGACATGTCAGCGGAAGATT 1659
DB 6724 GCGTTACGCGGTTTGTGTACCAACATTCACACGCTCTTTGACATGTCGCGGAGACT 6783
QY 1660 TTGAATGCAATATTTGCTGAACATTTCCACGCGGAGCCAGTATTTGAAACGACATCG 1719
DB 6784 TTGATGCAATATGACGAAACATTTCAAGAAAGTGAACCGGTACTGAGACGATATCG 6843
QY 1720 CGTGTGATTAAGCAAGCAAGCAAGCTATCGCCATTTGCGCGTTGATGATCTTGAAG 1779
DB 6844 CCGTGTGCAAAAAGCAAGCAAGCAGCTATGCGGTTAAACCGGCTGATGATCTTGAAG 6903
QY 1780 ACTTAGTGTGACCAACCGCTCTTAATTTGATGAGAGCGCGCTTCCGCAATATCAAT 1839
DB 6904 ACTGCGGTGAGCCACCACTACTCGACTTGATGATGAGTGGCCCTTTGAGAAATATCAT 6963
QY 1840 CTGACACTTACTACAGAAAGAGGTTTAATTTGTGTCATGATGAATCCGATATGT 1899
DB 6964 CCACCATCTGCGCCAGGCTACCGCTTCAATTTGGGCGATGATGAATCCGAAATGT 7023
QY 1900 TCTTAACGCTGTTGTCAACACATGATCAATATCATATTTGCTAGCAGATCTACGTC 1959
DB 7024 TCTTACGCTCTTTGTCAACAGATTTGAAATGCTTATGCGCAGAGATTTGAAG 7083
QY 1960 AACGTTAACCAAGCTCAGCGTCCGCGCTCTATGCGGAGATTAACATAGTCATGTG 2019
DB 7084 AGCGCTTAAACGTCCAAAATGTGACATTTATCGGAGCAACAACTTATACCGGAG 7143
QY 2020 TCGTCTCGACACCTGTATGCGGAGATGCGCACTTGTGCTGAACATGGAATGAATAA 2079
DB 7144 TAGTATCTGACAAAGAAATGCTGAGAGTGTGCCACTTGTCAACATGAGGTTAAGA 7203
QY 2080 TTATTTGATGATGATTTGATTTGATCAAAAGCACTTCTGTGGGGATTTATCTGTGTG 2139
DB 7204 TCATTTAGAGCATATCGGGAGAGAACCACTTACTTTGCGGTGATTCATCTTTGCAAG 7263
QY 2140 ACCAGATTAACAGGACAGCTTCAGAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
DB 7264 ATTCGTTACCTCCACAGCGTGTGCGTGGGAGACCCCTTAAAGGCTGTTAAAGTTGG 7323
QY 2200 GAAACCAATTCGCGAGTGAATATCCCAAGCTGGAACCGCGCGGCACTGATGATG 2259
DB 7324 GTAAACCGCTCCAGCGCAGCATGAGCAAGCAAGAAAGAGAGCCCTCTGCTAGATG 7383
QY 2260 AAGCAATGCAATGGAACAGAAATTTGAATTAAGCAAGTATGTAAGGCGCTGATATCA 2319
DB 7384 AAACAAAGCGGTGTTAGATAGATTAAGCAAGCACTTATGACATGCGCGGCAATC 7443
QY 2320 GATACGAGATCATCTGCGAGCGCTGATCATCGTCTGTCTGCACTTTAGCCGAAAGCG 2379
DB 7444 GGTATGAGTGAACAACATCACCTGTCTGTGCACTTGAATCTTTGCCAGAGCA 7503
QY 2380 TTAAGAAATTAAGAGCAATAAGAGAGCCCAATCACCTCTAGCGCTGACCTTAATAG 2439
DB 7504 AAAGGCAATTTCAAGCCATCAGAGGGAATTAAGCATCTCTACGCTGCTTAAATAGT 7563
QY 2440 TGACGTAGTAGA-----CAGCGCACTTACCCAGCGGAGATGTTTCA 2482
DB 7564 CAGCATATGATATTTCAATCTGATTAATACAAACACACACACATGATAGAGATTC 7623
QY 2483 TACCTTACGTTAACTTTCAACAGTTTACCTTCAAAATCCGATGCTTACCGAGATCCA 2542
DB 7624 TTAACATCTCGGCGCGCGCCCTTCCAGCGCCCACTGCAATGTGAGCGCGCGAGAA 7683
QY 2543 AACCTCTTAGCGCGCGTGTGAGCGCTTTGGCGCCCGCTGCTCTCAATGGAAT 2602
DB 7664 GAGGCAAGCGCGCCGAGTGTGCGCGCAATGAGGCTGCTTCCAAATCAAGCAACTGA 7743
QY 2603 CTTAGAGATGATAGTCAATCTTCAAAAC--GATCACTTAATCCGCGCGAG 2660

Db	7744	CCACAGCCGTCAGTGCCTTAAGTCACTTGACAGGCAACTAGACTCTCAAAACCCACGCCAC	7803
Oy	2661	GTCCACGCCCAAGAGAGAGAGTCTCTTAAGCAAACTTAAGCTTAAGTAAAAA	2720
Db	7804	GCCCCGCCCGCCGACAGAGAGAGGCGCCAAAGCAACCAACCGAAGCCGAAGAAACCA	7863
Oy	2721	AGAAAGCAAGCCAAAGAGACGAAAC---GAAAGCTTAAACAGGGAAACGACAACTGA	2777
Db	7864	AAACACAGAGAAAGAGAGAGAACTCTGCAAAACCCAAACCCGGAAGAGACAGGTA	7923
Oy	2778	TGTGATGAAGTTGAGAGTCGAGCAAGACTTTCGATC---ATGCTAAGCGCCAAATGA	2834
Db	7924	TGGCACTTAAAGTTGAGAGCCGACAGACTGTTCAAGCTCAAAAATGAGAGCGAGATGTCA	7983
Oy	2835	ATGATATATGCCCTGCGTTGTCCGAGAAAGCTGATGAACCACTCACGTTGAAGAAAA	2894
Db	7984	TCGGGACACGCACTGCGCCATGGAGAGAAAGTAAATGAACCACTCACTGTAAGAACTA	8043
Oy	2895	TTGATATGAGCAATTAGCGGCCGTGAATTGAAGAGGCTAGACATGATCGACTTGGAGT	2954
Db	8044	TTGACACACCTGTGCTATCAAAAGCTCAAAATTCACCAAGTCGTGACATACACATGAGT	8103
Oy	2955	ACGGGACGCTTCCCGAAGACATGAATCAGACGCTGACATGACACGACGACAAACAC	3014
Db	8104	TCGCACAGTTCGGGTCAACATGAGAGTGAAGGCTTCACTTAACACAGTGAACCTCTG	8163
Oy	3015	CGGGCTTCTCAACTGCGACCAACGCGCAGTCCAGTATGAGATGAGAGATTTTACCGTAC	3074
Db	8164	AAGGTTCTTACAATGGCACACCGAGCGGTGAGTATAGTGAAGGACGAGATTTACATCC	8223
Oy	3075	CGAAGAGATGGCGGGAAGGCGACAGCGGAAGCCGATCTTGACACACAGAGCGAG	3134
Db	8224	CCCCCGAGTGGAGGCGAGAGAGACAGTGTGCTGCATTTAGATTAATCAAGCCGCG	8283
Oy	3135	TTGTGGCATTTGTTCTAGAGGTGAAATGAGGGGACGCGTACGCGCTTCAATGGTCA	3194
Db	8284	TTGTGCGATATGCTCTCGAGGGGTGTATGAGGGAACAAGACCGCTTTTGCGTCTGA	8343
Oy	3195	CTTGAACCAACAAAGGGGTGACCAATTAGGATACCCCGAAGTTTCTGAACCGTGT---	3251
Db	8344	CCTGGAAATAGCAAAAGGGAAGACATCAAGACAAACCCGGAAAGGACAGAAAGTGTCTG	8403
Oy	3252	-----CACTAGTTACAGCGCTATGCGTCTTTGGAATGCAAGTTCATGCAAAAC	3305
Db	8404	CTGCACCACTGCTCAACGCGCACTGTGCTTGTGGAAGTGAAGTCCCATGCAATGCC	8463
Oy	3306	CACCCGTGCTATTTCACTGACGCGAAGACGAACCTGAAGTGTGSAABAAGAACGTGC	3365
Db	8464	CGCCCACTGCTACACCCGCGAACCATCAAGGCTCTGACATCTCTGAAGAGAACGTGA	8523
Oy	3366	ACAAATCCAAATTTACGACACGCTGTGAGGAAGCTTTGAAAT---GTGCATACGCGCGC	3422
Db	8524	ACCAAGAGCCTTAGACACACCTGCTCAACGCAATTTGCGGTGGGATGCTCCGCGAAG	8583
Oy	3423	CCAAACGAAGCATTTACGATGACTTCACTGACCAATGCTCCTACTGGGGTTCTGCCGT	3482
Db	8584	GTTAAAGAGGCTCACTAGCACTTTACCTTGACACGCGCGTACTTTGGGCAATGCTCGT	8643
Oy	3483	ATTGACAGCACTCAACGCGCGTGTTCACCCCAATTAATTAATGAAAGCTGGGACGAAT	3542
Db	8644	ACTGCACCAATATGAAACCGTGTCTTAAACCGATTAATATGAGAGGCTGGGATGAG	8703
Oy	3543	CTGATGATGAGATTCGATGAATCCAGGCTCGGCACAATTCGGGCTACAAATGAGCGACA	3602
Db	8704	CGAGAGCAACACCACTAGCAATCAGACTTCGCGCCAGTTTGGATACACAAACGGAG	8763
Oy	3603	CTGCGAGATGCACCAATTCGTTACATGTCTTTGACACGACCAATGACATCAAGAAG	3662
Db	8764	CAGCAAGCTCAAAATTAAGTACCGCTACATGTGCTGACAGAGATATATCTGTCAAAAG	8823
Oy	3663	ACAGTATGAGAAATTAAGCTATCAGCAATCTGACACCTGCCGTGTCTTGCCCAAAAG	3722
Db	8824	GCACCATGATGACATCAAGATCAGACACTTGAGGACCGTGTGAAGGCTTGTAGCTTCAAG	8883

Oy	3123	GGTACTTCCTGTTAAGCTCAATGTCTCTCCAGGTACAGTGTAAACCGTCAGTATCAACGACGC	3782
Db	8884	GATACCTTCTCTCCGCGAATGTCTCTCCAGGGACAGCGTAAAGGTTAAGCATAGCCAGTA	8943
Oy	3783	GAGCATCTGAGAAATTCATGACCGGTGGAGAAAAGATCAGAGAGAAAGTTGTGGTAGAG	3842
Db	8944	GCAACTAGACAGCGTCAATGCAAAATGGCCCGCAAGATTAACAAATTTGTTGGGACGGG	9003
Oy	3843	AGGAGTACTTGTTCACACCCGTCATGGAAGCTGGTAAAGTGCACGTTTACGATCACT	3902
Db	9004	AAAAATGTGACCTACCTCCCGTTTCCAGTAAAGAAATTCCTTGCACAGTGTACGACCGTC	9063
Oy	3903	TGAAGGAGAGGTGCGCGGGTACATTAACAATGACAGGCGAGGCCCAAGCGTATAGT	3962
Db	9064	TGAAGAAACAAACCGCCGGCTACATCACTATGCAAGGCCGGAGCCGCAATGCTATACAT	9123
Oy	3963	CCTATCTGAGGAGAGCGTCAGGCGAAGTGTACATTAACCACTTTGGCAAGACGTCA	4022
Db	9124	CCATCTGAGGAGATCATCAGGGAAGTTTAAACGGAAGCACACATCCGGAGAGAACATTA	9183
Oy	4023	CCATCCGAATGTAAATGTGGGACATACAGCACAGGTATCCGTGACGACCGCAACGAAGTGA	4082
Db	9184	CGTACGAGTCGAATGTGGGATTAACAAGACCGGAACCGTTACGACCCGTAACGAAATCA	9243
Oy	4083	ACGCGTCACTAAAGCAAAACAGTCACTTCTCTTACAGACGACCAACGAATGGGTCT	4142
Db	9244	CGGCTGTCAACCGCATCAAGCAAGTGCGTGCCTTAAGACGACCAACGAAGTGGTCT	9303
Oy	4143	TCAACTGCGCGGATCTTTATTAAGGACACAGACCACTAGTGCAGAGTAAATTGCACATTC	4202
Db	9304	TCAACTGCGCGGATCTCATGACACCGCCGACCAACAGGCGCAAGGGAAATTTGATTTGC	9363
Oy	4203	CATTCGCGCTTGACACCGACAGTCGTGCGCGGTCCGTAGTCTCACACGCTACAGTACGA	4262
Db	9364	CTTTCAAGCTGATCCGAGTACTGTCAATGTCCTGTGTGCCACGCGCCGAAAGTAAATAC	9423
Oy	4263	AGTGGTTCAAAAGGCATCAACCTTCACCTGACTGCAATGGAACCAACTTGTCTGCAACGA	4322
Db	9424	ACGCGTTTAAACATCATCAGCTCCCAATTAACACAGACCAATGTGACATTTGCTCACACCA	9483
Oy	4323	GAAATTTGGGGTGTGCGAGACAGACCAACAGAAATGGAATTAACGGGTCTACATCCAGA	4382
Db	9484	GGAACATPAGGGCAAAACCCGGAACCAACCACTGATGATGATATGCGAAACACGGTTAGAA	9543
Oy	4383	ATTTTCTGTGGGCGAGAAAGGCTGAGAGTACGTATGGGGTAAACATGAACCAATCAAG	4442
Db	9544	ACTTCAACGCTGACCGAGATGGCTGTGAATACATATGGGGCAATCAAGAACATGAAGGG	9603
Oy	4443	TCGTGGCCCAAGGATCGGACACAGGCGACCCACATGATGGCCCATGAGATCATCATCC	4502
Db	9604	TCATATGCCAAGATCTGCAACAGGAACCTTCACGATGGCCACAGAAATAGTACAGC	9663
Oy	4503	ACTATTTATCATCGGCATCTCAGTCTACACTGTGCATTTGTGTGTGTGCTGTCTCTA	4562
Db	9664	ATTACTATCATCGGCATCTCTGTGTACACCACTTATAGCCGTGCAATCAGCTGTGTGGGA	9723
Oy	4563	TCCTGTGAGGCACTGCATCATCAGCACTTGCAATGCGCAAGCAAGAGACATGCTCTGA	4622
Db	9724	TGAATGATGGGATTAACGTGTGACCACTTAATGTGCTGTAAAGCCGCGCTGATGAGCTCTGA	9783
Oy	4623	CGGCATATGCGCTTGACACCGAAGCAACGGTACCAACAGCATTAAGCGGTTTGTGCTGCA	4682
Db	9784	CGCAATATGCGCTGTGCCCCCAATATCCGTGATTCAACTTGTGCTGGCACTTTTGTGCTGTG	9843
Oy	4683	TTGCGGCCAACAGCTGTGAACATTTGGAGAACTTTGAACCATCTGTGTGTTTAAACAAC	4742
Db	9844	TTAGTGTGGCTAAATGCTGAACATTCACCGAGACCAATGATTTACTTAATGTGCGAACGCC	9903
Oy	4743	AACCGTTTCTGTGGACAGTTGTGCATCTCTGTGGACGCGCTTGTTATTTCTGTTCGGCT	4802
Db	9904	AGCGTCTTCTGTGGGTCCAGCTGTGTATACCTCTGGCGCGCTGTGTGTCTTAATGTGCTCT	9963

QY 4803 GCTTTTCATGCTGATGCTTTTATTTATTTGTTGAGGCGCTCTGCTGGGGAAGTAAACG 4862
 DB 9964 GTTGTCTATGCTGCTGCTCTTTTATTTAGTGTGGCGGCGCTACCTGGGGAAGTAAACG 10023
 QY 4883 CCTTGAACATGCGACCACTGTGCGCAATTTTCCGGGGAATCCCGTAAAGGCGTTGTCG 4922
 DB 10024 CTTAGGAACATGCGACCACTGTGCGCAATTTTCCAGATGCTGCTAAAGGCACTTTGTTG 10083
 QY 4923 AACGCGAGGTTTACGCGCACTTAACTGGAGATCAGCGGTGCTCATTCGGAATTAACAC 4982
 DB 10084 AAGGCGAGGTTTACGCGCACTTAACTGGAGATGCTCATTCGGAATTTTTCG 10143
 QY 4983 CTTCACTAACAGAGATGATGATGCTGCAATTTTCAACAGTCAATTCCTTCAACCAAG 5042
 DB 10144 CTTCACTAACAGAGATGATGATGCTGCAATTTTCAACAGTCAATTCCTTCAACCAAG 10203
 QY 5043 TTTAATGCTGGGGTCTCTGAGTGAAGGCAATCTCAAGGCGGATTAACATGCGCGG 5102
 DB 10204 TCAATGCTGGGGTCTCTGAGTGAAGGCAATCTCAAGGCGGATTAACATGCGCGG 10263
 QY 5103 TTTTGGGGGTTGATGCTTCAATTTGAGGAGGCGCAATGCTTCTGTCAGTGAAGA 5162
 DB 10264 TCTTTGGAGGGTGAACCTTCAATTTGAGGAGGCGCAATGCTTCTGTCAGTGAAGA 10323
 QY 5163 ACACAACTGATGATGAGGCTGATGATGCTGCTCAGACTGCACTATGATCAACGAG 5222
 DB 10324 ACACAACTGATGATGAGGCTGATGATGCTGCTCAGACTGCACTATGATCAACGAG 10383
 QY 5223 TCGCATAAAGTTTACACAGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5282
 DB 10384 AGGCATTAAGGCTGATGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10443
 QY 5283 CCACCGGCGACCTGATGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5342
 DB 10444 CTTACCATTTTCTTATGATGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10503
 QY 5343 AGGTATGATGAGGCGGCTGATGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5402
 DB 10504 AAGTATGATGAGGCGGCTGATGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10563
 QY 5403 GAAAGGCGCTTTTACAACTGATGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5462
 DB 10564 ATCCGCGCTGCTGATGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10623
 QY 5463 TCGGATGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5522
 DB 10624 TTTGAGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10683
 QY 5523 GGTGCTGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5582
 DB 10684 GGTGCTGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10743
 QY 5583 ATGAATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5642
 DB 10744 TCGAGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10803
 QY 5643 AAATGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5702
 DB 10804 AGATGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10863
 QY 5703 TTGATGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5762
 DB 10864 TTGATGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10923
 QY 5763 GCTGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5822
 DB 10924 AATGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 10983
 QY 5823 ACAAGGCTGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5882
 DB 10984 ATGATGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 11043
 QY 5883 AGGAAGGCTGATGATGCTGCTGCTGTAAGGCGGCTGCTGATGATTAACGCAACA 5942

DB 11044 AAGAGTCACAGTTTATGCTGAGAAAGGCGGCTGACAGTACCTTACAGCACCGCA 11103
 QY 5943 GCCCAACAGCAATTTTATGTTTGTGCTATGCGGAGAGAGTCCACTGCTGATGAT 6002
 DB 11104 GCCCAACAGCAATTTTATGTTTGTGCTATGCGGAGAGAGTCCACTGCTGATGAT 11163
 QY 6003 GTAAACAGCGGCGGACCACTTATGGAAGAACCAATTAAGTGCAGCAAGATTCAGG 6062
 DB 11164 GCATACAGCGGCGGACCACTTATGGAAGAACCAATTAAGTGCAGCAAGATTCAGG 11223
 QY 6063 CGGAGTTTCCAAACATCTTGTGAACCTGCTGCTTGTGCACTGTTTGGGAGGATATCC 6122
 DB 11224 CGGAGTTTCCAAACATCTTGTGAACCTGCTGCTTGTGCACTGTTTGGGAGGATATCC 11283
 QY 6123 TCAATGTTGAGGATTAATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6182
 DB 11284 TCAATGTTGAGGATTAATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11343
 QY 6183 GACTGAGGCGGAGACCTGA-CATAGCGGTAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6241
 DB 11344 GACTGAGGCGGAGACCTGA-CATAGCGGTAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11403
 QY 6242 GCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6268
 DB 11404 GCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11463
 QY 6269 AAAAGTATGATTTTCCGAGAGACAGTGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6318
 DB 11464 AAAAGTATGATTTTCCGAGAGACAGTGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11513

RESULT 10

BD071120

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PN

PD

PF

PI

C12N15/86

C12N15/33

C12N15/01

C12N15/10

A61K9/12

A61K8/00

CC

the bone marrow

genes in

CC

FT

CDS

CDS

Location/Qualifiers

1..1163

/organism="unidentified"

/mol type="genomic DNA"

/db xref="taxon:32644"

BASE COUNT

3291 a

3105 c

2907 g

2360 t

ORIGIN

Query Match

36.9%

Score 2396

DB 6

Length 11663

Best Local Similarity

64.6%

Pred. No. 0

Matches 3751

Conservative 0

Mismatches 1985

Indels 74

Gaps 9

OY	580	ACTGACGGTATGAAAGCCGGACCGTATATTTTCTCATTCGGAAACAGGCCAAGTCACTTC	633
Db	5707	ACTGCTTACCCGGGTAGTGGGTACATATTTTTCACGGAACAGGGCTTGGCACTTGC	5766
OY	640	AACAGAAATCAAGTACGTCAATGTGAACTTACAGAAACCTATATTTGGATCGGGCCGTCATG	699
Db	5767	AAAAGAAATCCGTTTCTGCAGAAACAGCTTACAGAAACGACCTTGGAGGCAATGTCTCG	5826
OY	700	AGAAATTTACGGCCCCCGCCCTCGATCTTCGAAAGAGAGAAAATGTTTACAGAAACCTGC	759
Db	5827	AAAGAAATCTACGCCCGCCGGTGCCTCGACAGTCGAAAGAGAAACAGCTCAAACTCAGTAAC	5886
OY	760	AATTATGCGCCCTCGAAGAAATATGAAACAGGTATCAATCAAGAAATAGAAAATATGA	819
Db	5887	AGATGATGCCCAACCGAAGCCAAACAAACAGGTACCACTCTCGAAAATAGAAAACCA	5946
OY	820	AAGCAATTTACAGCCGAGGAGCTCATTTTGTGATTTGGGACATATCTATCATCAGAAAGTA	879
Db	5947	AAGCCATATACCACTGAGGAGCTGTTTCAAGGGCTACGACTGATATTAATCTGCCACAG---	6003
OY	880	ATCCTGTGAGTGTTCACAGACTCATTTATCCTGTACCAATCTACTCTGTCCAACGGTATTA	939
Db	6004	ATCAGCCAGAAATGCTATTAAGATCACTACCCGAAACCATCGATTTCCAGAGTACCA	6063
OY	940	ACAGGTTTACACTCTGCAAGGTCGGGCTTAAACGTGCACCTTATGTTATCCAGAGAAAT	999
Db	6064	CGAATCACTCTGACCCCAAAAGTTTGTGTAGCTTTGTAAACAATCTCTGATAGAAAT	6123
OY	1000	ACCTTACAGTACGCCAGTTATTTATTAACAGATGAATAAGATGCGTATCTTGACATGTGG	1059
Db	6124	ACCCGACGGTAGCATCTTATCAGATCACCGACGAGTACGATGCTTACTTGTGATATGTGAG	6183
OY	1060	ACGGGCACTCGTGTCTAGTACAGCCACTTTTGTTCGGGCTTAACTGAGAAAGCTACC	1119
Db	6184	ACGGGACAGTGGCTTGCTTGAATCTGAACTTTTGTGCCCGGCAAGCTTAAAGTTTACC	6243
OY	1120	CAAGAGACATAGCTATTTTGCAGCCAGATTAAGTACGCCGTCCATCGCTTATACGA	1179
Db	6244	CGAAAGACACGAGATATAGAGCCCAAAACATCCGACGTGGGTTTCATACACGATGCAGA	6303
OY	1180	ATATCACTTCAAAATGTATTTGGTGTGAGCTACTTAAAGAAATTTGCAAGCTTACCCAAATGC	1239
Db	6304	ACAGCTTCCAAAACGTGTCTCATTTGCGGACTTAAAGAAACCTGCACATCTCACAAATGC	6363
OY	1240	GAGAAATTACTGTCTTATGATTCGGCGGCAATTTATGTTGATTGTTTCAAGAAATACGAT	1299
Db	6364	GTGAATCTGCCAAACACTGGACTGTAGGAACTTCAACGTTGAATGCTTTCGAAAATATGAT	6423
OY	1300	GCAATGATGAGTACTTGGGATACCTTTCCGATTAACCTTATTCGGCTAACTACAGAAACG	1359
Db	6424	GCAATGAGAGATATTTGGAGAGATTTTCCGAAACCAATTAAGATATCTACTGATTCG	6483
OY	1360	TTACGCAATATGTGCAAGAGCTGAAAGGGCCGAAAGCAGCAGCATTTGTTGCAATACTC	1419
Db	6484	TTACCGCATAGGTGGCCAGACTGAAAGGCCCTTAAGCCCGCGCACTGTTCGCAAAAGACG	6543
OY	1420	ATTAATCTTAAACCGTTGACAGATATCCATATGATCAATTCGTCAATGATCTTAAAGAG	1479
Db	6544	ATTAATTTGGTCCCATTTGCAAGAAAGCTTATGAGATAGATTTGTCATGACATGAAAAG	6603
OY	1480	ATGTCAAAGTTACTCCCGGACAGAAACATTAACAGAGAACCGGCTTAAGGTGACAGTTATTC	1539
Db	6604	ACGTAAAGTTTACACTGGCACAGAAACACACAGAAAGAAACCGAAATTAACAAGTATAC	6663
OY	1540	AGGCTGCAGATCCCTTGTCTACCGCTTACTTTTGGGGATCCATCGGAATTAATGTCGTA	1599
Db	6664	AAGCCGAGAAACCCCTGGGACCGCTTACTATGCGGATCCACCGGAGTTAATGTCGCA	6723
OY	1600	GACTGAATGCGGTGCTTTGCCAAATATTCATATCTCTTGGACATGTTCAGCGGAAGATT	1658
Db	6724	GGCTTACAGCCGTTTGTCTACCCACATTCACACGCTTTTGGACATGTTCGCGGAGGACT	6783

Qy	1660	TTGATGCGAATATATGCTGAAACATTTTCCACACGCGCGACCCAGATATTTGGAAACGGACATCG	1712
Db	6784	TTGAATGCAATATATGACGAAGAACATTTCAAGCAAGGTGACCCGGTACTGAGACGGATATCG	6843
Qy	1720	CGTGGTTTGAATAAAGCGAAGACGACGCTATCGGCATTTTCGGCGCTTGATGATCCCTTAGG	1779
Db	6844	CCTGGTTTGAACAAAGCCGACACGCTATAGGGGTTTAAACCGGCTGATGATCTTTGGAG	6903
Qy	1780	ACTTAGGTGTGACCAACCGCTCTTAAATTTGATAGAGCGGCGCTTCGGCAATATCAAT	1839
Db	6904	ACCTGGGTGTGACCAACCACTTACCTGACCTTGATCGAGTGGCGCTTTGGAGAAATATCAT	6963
Qy	1900	TCTTAAAGCGTGTGTTGTCAACACATGACATATCATGATGTTGTAGCAGAGATACATAAGT	1959
Db	7024	TCTCAGCGCTTGTGTCAACACAGTTCTGAATGTGTTATGTGCGACGAGATATTTGGAG	7083
Qy	1960	AACGGTTAACACGTCAGCGTGCAGCGCTCTATCGCGAGATTAACATATGTCATGCTG	2019
Db	7084	AGCGGCTTAAACGTCGAATGTGACAGATTTATCGGAGAGACAACTTATACAGGAG	7143
Qy	2020	TGCTGTCCGACACCTTGATGTGCGGAGAGATGCCGCACTTGGCTGAACATGGAATAAAA	2079
Db	7144	TAGTATCTGCAAGAAAGAAATGGCTAGAGGTGTGCACCTGGCTCAACATGAGGTTAAGA	7203
Qy	2080	TTATTTGATGAGATTATTTGGTATCAAGACCCCTACTTCTGTGCGGGGATTTATCTGTGG	2139
Db	7204	TCATTTGACGAGTCAATCGGCGAGAGACCACTTACTTCTGGGGTGATTTCACTCTTGACAG	7263
Qy	2140	ACCAATTAACAGGACAGCAGCTGCAGAGTCCGACAGCCCTTAAAAAGCTTTTAAAGCTTG	2199
Db	7264	ATTGGTTACTTCCACAGCGTGTGCGGTGGGAGACCCCTTAAAAAGCTGTTTAAAGTTGG	7323
Qy	2200	GAAAAACCATTTGCCAGTGCATATATACCAAGACTGCGACCGCGCGGCACTGCATGATG	2259
Db	7324	GTAACCCCTCCACGCGCAGCATAGACAGAACGAAAGACGAGAACGCGCTCTGTTAGATG	7383
Qy	2260	AAGCAATGCGATGGAACGAATTTGGAATATACGACGAGTTGATGTAAGGCGGTAGAAATCCA	2319
Db	7384	AAACAAAGGCGTGTTTATGATGATGATATACAGACACTTATGCACTGGCGGTGGCAATC	7443
Qy	2320	GATACGAGATCATACTGCGAGGCGCTGCATCATCAGTCTCTGTCCACGTTAGCCGAAACGG	2379
Db	7444	GGTATGAGGTGACAAACATACACACTGTCTGTGGGATTTAGAACTTTTGGCCAGAGCA	7503
Qy	2380	TTAAGACCTTCAAGGACATTAAGGAGAGCCCAATCACCTTATAGGCTGATCTTAATATGG	2439
Db	7504	AAAGGCACTTCAAGCCATCAGAGGGGAAATAAAGCATCTTCAAGGTGTCTTAAATAGT	7563
Qy	2440	TGACCTAATATA-----CAGCGACCTAACCCAGCGAGAACTTTTCCA	2482
Db	7564	CAGCATATATACATTTTCACTGACTAATACCAACACACACATGATATGAGGATTTCT	7623
Qy	2483	TACCTCAGCTGAACCTTTCACACGATTTTACCTTCAAAATCCGATGGCTTACCGAGATCCA	2542
Db	7624	TTTAACATGCTGGCGCGCGCCCTTTCCAGGCCCACTGCCCATGTGGAGGCGCGGAGAA	7683
Qy	2543	AAACCTCTTAGAGCGCGCTGAGAGCGGTTTCGCGCCCGCGTGTCTCAATATGAAAGAT	2602
Db	7684	GGAAGGCAAGGCGCGCCGATGCGCTTCGCGCAATGGGCTGCTTCCCAATCCAGACACTGA	7743
Qy	2603	CTTAGGAGTGCATATGATCACTTTCATTTCAAAAC--GATCACTAATTCGCGCGCAG	2660
Db	7744	CCACAGCGGTAGTGCCCTATGATCTTGGACAGGCAACTGACCTTCAAAACCCCAAGCCAC	7803
Qy	2661	GTCCACCGCCAAAGAAAGAAAGATGTGCTCTTAAGCCAAACCTACTCAGCTTAAAAAGA	2720
Db	7804	GCCGCGCGCGCGCCAGAAAGACAGGCGCCAAAGCAACACGAAAGCCGAAGCAAGAAACCA	7863
Qy	2721	AGAACGACGAAGCAAGAGCAGAAC--GCAAGCTTAAACAGGGAACGACCAAGCTA	2777

Db	7864	AAACACAGAGAGAGAGAGAGAACCACTGCAAAACCCAAACCCGGAAAGACAGCGTGA	79223
Oy	2778	TGTGTATCAATTTGGAGTTCGGACAAAGACATTTCCGATC---ATGCTGAACGGCCCAAGTGA	28348
Db	7924	TGGCACTTAAGTTGGAGGCCGACAGACTGTTCGAGTCCAAAATAGAGACGGAGATGTCA	79878
Oy	2835	ATGATATATGCTTCGCTGTTCGGAGGAAGGCTGATGAACCACTCCACGTTGAAGAAAA	28948
Db	7984	TCGGGCAAGCACTGGCCATGTGAGAGAAAGATATGAACCACTCCACGTTGAAGAAAA	80434
Oy	2895	TTGATTAATGAGCAATTAGCGGCCGTGAATTTGAAGAAGGCTAGCATGTATCCACTTTGAGT	29558
Db	8044	TTGACACACCTGTGCTATCAAGCTCAAAATTTCCAAAGTCTGACATATACAGATGAGAGT	81038
Oy	2955	ACGGGACGTTTCCCGAGAACATGAATAGACACGCTCGAGTACACGCGCAAAACAC	30148
Db	8104	TCGCACACTTCCCGGTCAACATGAGAAAGTGAAGGCTTACCTCAACAGTGAACCTCTG	81638
Oy	3015	CGGGCTTTCAACCTGGACCAACCGGCGAGTCCAGTATGAGAAATGGAATTTACCCGTAC	30748
Db	8164	AAGGTTTCTAACACTGGCACACGAGCGGTCCAGTATATGAGAGGACAGATTTATCATTC	82228
Oy	3075	CGAGAGAGTGGGCGGGAAAGGCCACACCGGAACCGATCTCTGACAAACAGAGGACAG	31348
Db	8224	CCCGGGAGTGGAGGACAGAGAGACAGATGTGCTCCGATTAATGATTAATCAGGCCCGG	82838
Oy	3135	TTGTGCTATTGTTTCTTAAGAGGTGCAAAATGAGGGACCGGTACCGCGTTTCAGTGCTCA	31948
Db	8284	TTGTTCGCAATGTCTCTGGAGGGGCTGATGAGGAAACAAAGACCGCTTTTCGGTGTCA	83438
Oy	3195	CTTGGAACAGAAAGGGGTGAACCATTAAGGATACCCCGGAAGTTCTGAACCGTGT---	32518
Db	8344	CCTGGAAATAGCAAAAGGAAACAAATCAAGACAAACCCCGAAGGACAGAAAGTGTGTCTG	84038
Oy	3252	-----CACTAGTTTACAGCGCTATGCGGTCTTTGCAATGTCAAGTTCCCATGCGCAAAAC	33058
Db	8404	CTGCACCACTGCTACCGGCAATGTGCTTGTCTTGAAGACGTACGCTTCCCATGCAATCGCC	84638
Oy	3306	CACCCGTGTGCTATTCACTGAACGCCAGAAACGAAACCTGACGCTGTGAAGAGAACGTCG	33658
Db	8464	CGCCCACTGTTCACCCCGGCAACCATCAAGACTCTGACATCTCTGAAGAGAACGTCGA	85238
Oy	3366	ACAAATCCAAATTAACGACACGCTGTCTGAGAAAGCTCTTGAAAT---GTCCATACGCGCCGC	34228
Db	8524	ACCAGAGCGCTTAGACACACCTGCTCAACGCCATATTTGGGTCGGAGTGTGTCCGCAAA	85838
Oy	3423	CCAAACGAAGCATTTACGATGACTTTCACACTGACCAATGCTCTACCTGGGGTTCTGCCGT	34828
Db	8584	GTAAGAAGAGGCTCACTACGACATTTTACCTTGAACAGCCCGTACTTGGGCAATGCTGT	86438
Oy	3483	ATTGAGACACTCAACGCGGTGTTTCAAGCCCAATTAATTAATGAAACCTGTGGAGCAAT	35428
Db	8644	ACTGTACCAATACTGAACCGTGTCTTTAACCCGATTAAGATCGAGAGGTCTGGAGTGAAG	87038
Oy	3543	CTGATGATGATTCGATTTGAATTCAGAGTCTCGGCAATTTGGGCTACATACAGGACGCA	36028
Db	8704	CGAGAGCAACACCATTAAGCATACAGATTTCCGCCAATTTGATAGCACCAAAGCGAG	87638
Oy	3603	CTGCGGATGTCACCAAATTCGTTTACATGTCTTTTGCACCAAGACCATGACATCAAGAG	36628
Db	8764	CAGCAAGCTCAAAATAGTACCGCTTCAATGTGCTGACAGAGATCAATCTGTCAAAAGAG	88238
Oy	3663	ACAGATGAGAAAAATAGCTATTCAGCAATCTGAGACCGTGCCTGTCTTGGCCCAAGAG	37228
Db	8824	GCAACCATGATGACATCAAGATCAAGCACTTCAAGACCGGTGTGAAGGCTTATGCTTCAAG	88838
Oy	3723	GGTACTTCTGTATGCTCAATGTCTCCAGGTGACAGTGAACCGTCACTATTCACAGACG	37828
Db	8884	GATACTTCTCTCTCGCAAGTGTCTCTCAAGGGGACAGGTAAACGTTTACATAGAGACTA	89438
Oy	3783	GAGCATCTGAAGATTCATGACACCGTGAAGAAAAAGATCAGAGAGAGTTTGTCCGTAGAG	38428

Db	8944	GCAACTACGACAGCTATGCACAATGGCCCGCAAGATTAACAAATTCGTGGACGGG	9003
OY	3843	AGGAGTACTTGTTCCACCCGTCATAGAAAGCTGGTAAAGTGCACGTTTACGATCACT	3902
Db	9004	AAAAATATGACCTACCTCCCGTTCAAGGTAAAGAAATTCCTTGACAGTGTACGCCGTC	9063
OY	3903	TGAAGAGACGTCTGCCGGGTACATTAACAATGACAGGCCAGGCCACACCGCTATPACT	3962
Db	9064	TGAAGAAACACCCGCGCTACATCATATGACAGGCCGGGACCGCATGCTTATACAT	9123
OY	3963	CCATCTGAGGAAAGCCTCAGGGAGTGTACATTAACCACTTGTGCAAGAACTCA	4022
Db	9124	CTATCTGAGGAAATCATCGGAAAGTTTACCGAAGCCACCATTCGGGAAACATTA	9183
OY	4023	CCTACGATATTAAGTGTGCGATACAGACAGAGTATCTGTGACACGCGAACGAAGATGA	4082
Db	9184	CGTACGAGTCCAAAGTGGGGGATTTACAAAGCCGGAAACGTTACGACCCGTTACGAAATCA	9243
OY	4083	ACGGCTGCACTAAAGCAAAACAGTGCATTGCTTACAAAGACGACCAACGAAATGGGTCT	4142
Db	9244	CGGGCTGACCGGCATCAAGCAATGGCGTCTAATGAAGCGACCAACGAAGTGGTCT	9303
OY	4143	TCGACTGCGCGGATCTTTATAGGACACAGAACCACTCAGTGCAGGTAATTTGCACTTTC	4202
Db	9304	TCAACTGCGCGGATCTGCATCAGACAGCGCGACCAACGCGCCAAAGGAAATTTGCATTGTC	9363
OY	4203	CATTCCGCTTGACACGACAGTCTGCGCCGGTTCCGTTAGCTGCACAGCCTTACAGTACGA	4262
Db	9364	CTTTCAAGCTGATCCCGAGTACCTGTGATGTCTCTTGTGCCACGCGCCGAACGTATGTC	9423
OY	4263	AGTGTTCAAAGGCATCACCTTCACCTGATCTGCAATGGCACCAATTTCTGCACACGA	4322
Db	9424	ACGGCTTTAAACATCATCAGCCTCCAATTATGACACAGACCATGTGACATTGTCTACACCA	9483
OY	4323	GAAAAATTTGGGGCTGCGAGACAGACCGAACAGACAGATATGAGGTCTACATCCAGCA	4382
Db	9484	GGAACATGAGGGGCAAAACCGGAAACCAACACTGAATGATCATCGGAAACACGGTTAGAA	9543
OY	4383	ATTTTCTGTGGGGCGAAGAGGCTGAGATGCTATGGGGTAAACAATGAACAGTACAGAG	4442
Db	9544	ACTTACACCGTGCAGCCGAGATGCGCTGGAATACATATGGGGCAATACGAAACAGTAAAGG	9603
OY	4443	TTGGGGCCGAGAGTGGGACCAAGGCGACCCACATGSAATGGCGCATGAGATCATCATCC	4502
Db	9604	TCTATGCCAAAGTGTGCACAGAGAGACCTCAGCATGGCCACACGAATATGATCACAC	9663
OY	4503	ACTATTTATCATCGGCATCCAGTGTACATCTGTCACTTGTGCTGTGTGCTCTTGTCTTA	4562
Db	9664	ATTACTTATCATCGGCATCTCTGTGTACACATCTTTACCGCTGCATCAGCTGTGGCGCA	9723
OY	4563	TCCGTGATGGGACCTGCATATCATAGACAGTTGCAATGGCCAAAGCAAGAAAGACCTGCTGA	4622
Db	9724	TGATGATTTGGCGTAACTGTTGCGACATTAATGTGCTGTAAAGCGCCCTGTAGTGTCTTA	9783
OY	4623	CGGCATACGGGCTTTGACCGGAACGCAACGGTATCCACAGCATTAAGCGGTTTTGTGCTGCA	4682
Db	9784	CGCATATATGCGCTGTGGGCCCAAAAGCGGTATTTCAACTTGTGCGGACATTTTGTGTGTG	9843
OY	4683	TTGGGCCCAACACGCTGAAACATTTGGAGAAACTTTGAACCACTGTGTGTTTAACACCC	4742
Db	9844	TTAGTGTGGCTTAATGTGTGAACATTCACGAGACCAATGATTACTTATGTGTGAACAGCC	9903
OY	4743	AAACGTTTCTGTGGGACAGTTGTGCAATTCCTGTGACAGCGGTGTGTAATTCGTGCCGT	4802
Db	9904	AGCGTCTTCTGTGGGTTCAGCTGTGTATACCTTGTGCGGTGTGTGCTTCTTAATGCGCT	9963
OY	4803	GCTTTTCATGCTGACATGCTTTTATTAATGGTTGACAGCGCTGTGCTGTGAGAAAGTATAGC	4862
Db	9964	GTTGTCTATGCTGCTGCGCTTTTATAGTGGTTGCGCGGCCCTTAACCTGGGAAAGTATAGC	10023
OY	4863	CCTTCGAACATGCGACCACTGTGCCAAATGTTCCGGGGGATCCCGTATTAAGCGCTTGTGTCG	4922
Db	10024	CTTACGAACATGTGACCACTGTTCCAAATGTGCAACAGATATCGTATTAAGCACTTGTGTG	10083

QY 4923 AACGCGAGTTAGCGCGCACTTAACCTGAGATCAGGTCGTCTCATCGGAATTAAAC 4982
 DB 10084 AAGGAGAGGGTACGCCCGCTCAATTGGAGATTACTGTCTATCTCTCGAGGTTTGC 10143
 QY 4983 CTTCACATTAACAGAGTACGTGACCTGCAAAATTCACACAGTATTCCTTACACACAG 5042
 DB 10144 CTTCACCAACAGAGTACCTTACCTGCAAAATTCACACAGTATTCCTTACACACAG 10203
 QY 5043 TTAAATCTGCGGCTCCCTGAGTGCAGAGCATCTCAAGAGCGGATTTACATGCGCG 5102
 DB 10204 TCAGATCTGCGGCTCCCTGAGTGCAGAGCATCTCAAGAGCGGATTTACATGCGCG 10263
 QY 5103 TTTTTCGCGGTGTACCTCTTCAATGTGAGGAGCGCAATATGCTCTGTGACAGTGA 5162
 DB 10264 TCTTTGAGGGGTATACCTCTTCAATGTGAGGAGCGCAATATGCTCTGTGACAGTGA 10323
 QY 5163 ACACACAAGTGTAGGAGCGGTACGTGAGTTGCTCCAGACTGCACTATAGATCAAGCAG 5222
 DB 10324 ACAGCCAGATGTAGTGAAGCGGTACGTGAGTTGCTCCAGACTGCACTATAGATCAAGCAG 10383
 QY 5223 TCGGACTAAAAGTTCAACAGCTGTCTGAAAAGTCCGCTGCTGATATGATACGGCAACA 5282
 DB 10384 AGGGGATTAAGGTGCATATGCTCCGCAATGAAAGTGAAGTGCATATGATACGGCAACA 10443
 QY 5283 CCACCGCGCACTGTGATACGTTTGTCAATGCGGTCAAGCGGAGTTCCCTCAAGGAGCTGA 5342
 DB 10444 CTACCACTTTCTTGATGTGTAGCTGAAAGGAGTCAACCGAGAACTCTTAAAGACTGA 10503
 QY 5343 AGGTATAGCAGGCGCGATATCAGCCGCTTTTTCACCTTTGACATTAAGTCTGATCA 5402
 DB 10504 AAGTCATAGCTGAGCAATTCAGCATTTGTTTACACCTTGATCAACAGGTCGTTATCA 10563
 QY 5403 GAAAGGCGCTTTTACATCAGACTTCCCTGAGTATGAGCTATGAAACAGAGGCT 5462
 DB 10564 ATCCGCGCTGTGTACAACTATGACTTTTCGGAATGCGAGCATGAAACAGAGGCT 10623
 QY 5463 TCGGCGATATCAAGCATCTCGCTTGATGTACAGCATATGATAGCCGCGCTGACATAC 5522
 DB 10624 TTGAGACATTCAGGCTACCTCTGACTACGAAAGACTCATGCGCGACAGACATTA 10683
 QY 5523 GGTGCTGAAGCTTCTGTCAAGAACATCAAGTCCCTCAACCCCAAGCAGTATCAGGT 5582
 DB 10684 GGTACTCAAGCTTCCGCGCAAGAACGTCATGTCCTGACACGAGCGCCGCTCTGGAT 10743
 QY 5583 ATGAAATGTGAAGAACAACCTCAGAGCAACCCCTGCAAGAAACAGCAGCATTTGGATTA 5642
 DB 10744 TCGAGATGTGAAAAACAACCTCAGAGCGCCGCACTGCAAGAAACCGCCCTTTTGGTGA 10803
 QY 5643 AAATTGAAGTGAAGCTTGTGAGAGGTACTAAGTGTGAGGCAATCCCTATCTGCA 5702
 DB 10804 AGATTGCAATTAATCCGCTTGTGAGAGGTGAGTGTCTATACGGAACTTCCCATTTCTA 10863
 QY 5703 TTGACATCCCTGATGCACTTTTGTGAGATCATGATCAACCAATTTTGAAGTTA 5762
 DB 10864 TTGACATCCCGAAGCTGCTTTTATCAGGACATGATGACCACTGCTTCAACAGTCA 10923
 QY 5763 GCTGCACTGACGAGTGTCTATTTTGTGAGACTTTGTGTGCTTCTTAACTATACGT 5822
 DB 10924 AATGTGATGTCAAGTGTGACCTTATTCAGGCGGATTCGAGGAGTGTGCTGACGT 10983
 QY 5823 ACAAGGCTGACAGGAGGAGGATTCAGTTCAGTTCCTCCACTCCACGACAGCTGTTTGA 5882
 DB 10984 ATGTATCCGACCGGAGAGGAAATGCTTGTATCTTCCATTTCCGACAGACAGCCTTCC 11043
 QY 5883 AGGAGCGACCAACATGTGACTGCGGTAGGACAGCATTAATCATTTTGAACATGCA 5942
 DB 11044 AAGGTGACAGTTCATGCTCGGAGAAAGAGCGGTGACAGTACATTTGACACCGGA 11103
 QY 5943 GCCCAAGCAAAATTTTATGTTGCTATGCGGCAAGAGTCCACTGCAATGTGAT 6002
 DB 11104 GCCCAAGCGCAAACTTATGTTATGCTGTGTGTGAAGAAACAATGCAATGACAGAT 11163

QY 6003 GTAAACCAACCGCGCAACCATATATTGAGAACCAATAAATGACCAAGAAATTCAGG 6062
 DB 11164 GCAAAACCAACGCGTATCATATGATGAGCAACCCCGACCAAAATATGACCAAGAAATTCAG 11223
 QY 6063 CGGCACTTCCAAACATCTTTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6122
 DB 11224 CCGCATCTCAAAAACCTTCAATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11283
 QY 6123 TCATGTTGAGACTTATATGATGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6182
 DB 11284 TATTAATTAATGAGACTTATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11343
 QY 6183 GACTGACCGCGCACTGA-CATAGCGTAAACTGATGATCTTCCAGAAAGCTGT 6241
 DB 11344 GACCGTACCGCCCAATGACCCGACGCAAACTGATGATCTTCCAGAAAGCTGT 11403
 QY 6242 GCATATGCAACGCGCGCT-----TGACACT 6288
 DB 11404 GCATATGCAATCAGCGCTGTATATGATCCCGCTTACCGCGGCAATATAGCAAC 11463
 QY 6288 AAAAATGATATTTCCGAGGAGCAAGTGAATATGCTGTCAGTGT 6318
 DB 11464 AAAAATGATATTTCCGAGGAGCAAGTGAATATGCTGTCAGTGT 11513
 RESULT 11
 146902
 LOCUS 146902 11663 bp DNA linear PAT 07-OCT-1997
 DEFINITION Sequence 1 from patent US 5639650.
 ACCESSION 146902
 VERSION 146902.1 GI:2470867
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 11663)
 AUTHORS Johnston, R.E., Davis, N.L. and Simpson, D.A.
 TITLE DNA clone for South African Arbovirus No. 86
 JOURNAL Patent: US 5639650-A 1 17-JUN-1997;
 FEATURES
 source 1..11663 /organism="unknown"
 BASE COUNT 3291 a 3105 c 2907 g 2360 t
 ORIGIN
 Query Match 36.9%; Score 2396; DB 6; Length 11663;
 Best Local Similarity 64.6%; Pred. No. 0; Mismatches 1985; Indels 74; Gaps 9;
 Matches 3751; Conservative 0;

QY 580 ACTGACGGTATGAAGCGGAGCGTATTTTCTCATCGAAACAGCCAAAGTCACTTC 639
 DB 5707 ACTGCTAACCGGGGTAGTGGTACATATTTTGAAGACAGCGCCCTGGGCACTTGC 5766
 QY 640 AACGAAATCAGTACGTCATATGTAATCAAGAACTTATTTGATCGGGCGCTCATG 699
 DB 5767 AAAAGAGTCCGTTCTCAGAACCAAGCTTACAGAACCGACCTTGAACGCAATTTCTGG 5826
 QY 700 AGAAGTATTAAGCCCGCGCTGATCTCGAAAGAGGAAATTTTACAGAAAGAACTGC 759
 DB 5827 AAAGAAATCTAGCCCGCTGCTGACAGTGAAGAGGAAACAGCTCAAACTCGGTACC 5886
 QY 760 AATATGCGGCTCTGAAGAAATGAGAGCAGTATCATACGAAAGTGAATAATATGA 819
 DB 5887 AGATGATGCCCAACGAGCAACAAAGAGTACAGTCTCGAAAGTGAATAATATGA 5946
 QY 820 AAGCAATTAAGCGGAGCGACTATTTCTGATTTGGACATATCTATCATCAGAGTGA 879
 DB 5947 AAGCATTAACCATGACGAGTCTTCAAGGCTTACAGAGTATTAATCTGCGCAG--- 6003
 QY 880 ATCTGTGAGTGTTAAGAGTCAATATCTGTGACCAATCTACTGTCACAGGTATTA 939
 DB 6004 ATCAGCAGATGCTATTAAGATCACTTACCCGAAACATGATATTCAGCAGGTATACAG 6063

Qy 940 ACAGTTTACATCTGACAGAGTCCGGTTAAACGTGCACTTACTTATTCAGAGAAATT 999
Db 6064 CGAACTACTCTGACCAAGTTTGTCTGATAGCTTTGTAACAACATCTGCAATGAGATT 6123
Qy 1000 ACCCTACGTCAGCCAGTTATTGTATTAACAGATGAATACGATGCGTATCTTGACATGTGG 1059
Db 6124 ACCCGACGTCAGCTTATACATGATCACCAGACGATGATGCTTACTTGGATATGTAG 6183
Qy 1060 ACGGCGCATCGTGTCTAGATACAGCCATTTTGTCCGCTTAACGTGAACTACC 1119
Db 6184 ACGGACGTCGCTGTCTAGTACTGAACTTTTGTCCGCTTAACGTGAACTACC 6243
Qy 1120 CAAGAAGCATAGCTATTGTGACGAGATMAATCAGCGTCCCATCGCTATACAGA 1179
Db 6244 CGAAAAGACGAGATATAGAGCCCAAAATCCGAGTGGCTTCATACAGAGTACAGA 6303
Qy 1180 ATACATTCAAAATGTATGTGCTGACGCTACTAAAGAAATGGCAAGTTACCAATGC 1239
Db 6304 ACAGCTTGCAAAACCTGCTCATGTGCGGACTAAAGAACTGCACACCAAAATGC 6363
Qy 1240 GAGATTACCTGTCTTATATCGGGGCAATTAATGTTATGTTTCAAGAAATGCGAT 1299
Db 6364 GTGAATCTCCAACTAGCTACAGCAATTCACGTTGAATGCTTGAAAATATGCAAT 6423
Qy 1300 GCAATGATGATGATGAGTACCTTTCCGATTAACCTATTCCGCTAATCTACAGAAAG 1359
Db 6424 GCATGACGATATTTGGAGAGAGTTTGCCTCGAAACCAATTAAGATCACTATGAGTTGC 6483
Qy 1360 TTAAGCAATATGTGCAAAAGCTGAAAGGCGGAAAGACAGCATTTGTTGCAATATCT 1419
Db 6484 TTACCGCATATGTCGCGACAGCTGAAAGGCGCTTAAGGCGCGCACTGTTCCGAAAGACG 6543
Qy 1420 AATAATCTAAAACCGTTGAGAGAGATACCAATGATCAATTTGTCATGATCTTAAAGAG 1479
Db 6544 AATAATTTGCTCCATTTGCAAGAAAGTCTATGATGATGATGCTATGCAATGAAAGAG 6603
Qy 1480 ATGTCAAAATTAATCTCCGCAAGAAACATATACAGAGAGCGGCTTAAGGTGACGTTATTC 1539
Db 6604 ACGTGAAGATTAACATCTGCGCAAGAAACACAGAAAGAAAGACCAAAATTAACATGATATC 6663
Qy 1540 AGGCTGAGATCCCTTGTCTAACCGCTTACCTTTGCGGAGTCCATGCGGAATTAATCTC 1599
Db 6664 AAGCGCGCAACCCCTGCGGACCGGCTTACTATGCGGAGTCCACCGGAGATTAAGTCCCA 6723
Qy 1600 GACTGAATCGCGGTCTTGTGCAAAATATCCATCTCTTGTGCACTGTGACGCGGAATTT 1659
Db 6724 GGGTTACGCGGTTTGTCTACCAACATTTCAACGCTCTTGTGCAATGTGCGGAGGACT 6783
Qy 1660 TTGATGCAATTAATGCTGAACATTTCCACACGCGGACCCAGTATTTGAAAACGAGATCG 1719
Db 6784 TTGATGCAATCAATAGACAGAACATTTCAAGAGTGAACCGGATCTGGAACGAGATATCG 6843
Qy 1720 CGTGTGATTAAGCGAAGACGACGCTATCGCCATTTGCGCGTTGATGATCTCTTGAAG 1779
Db 6844 CCTGTGTGACAAAAGCCAGACGACGCTATGCGCTTAACCGGCTGATGATCTTTGGAAG 6903
Qy 1780 ACTTATGATGATGACCAACCGCTCTTGAATTTGATGAGAGCGGCTTGGCAATATCAT 1839
Db 6904 ACCGTGGGTGAGCAACACACTATCTGACTGATGATGCGCTTTGGAAGAAATATCAT 6963
Qy 1840 CTGTGACCTTACTACAGAGAGGTTTAAATTTGTGCCATGATGAATCCGTTATGT 1899
Db 6964 CCACCATCTGCGCAAGGTTACCGTTTCAATTTGGGGCGATGATGAATCCGAAATGT 7023
Qy 1900 TCTTAACGCTGTTGTCAACAACATTAATCATGATTTGCTAGAGATCTACATCTG 1959
Db 7024 TCTTCAAGCTCTTGTCAACAACATTTGAATGTGTTATGCGCCAGAGATATGAGG 7083
Qy 1960 AAGCGTTAACACGTCAGCGTGCAGGCTCTATGTGCGAGATTAACATAGTGCATG 2019
Db 7084 AAGCGTTTAAACGTCAAATATGTGCAATTTATCGGAGACGAAACATTTATACAGGAG 7143
Qy 2020 TCGTCTCCGACACTTATGCGGAGAGATGCGCACTTGTGTAACATGAAATGTAATA 2079

Db 7144 TAGATCTGACAAAGAAATGCTGAGAGTGTGCACTGCTCAACATGAGGTTAAGA 7203
Qy 2080 TTATTTGATGCACTTATTTGTATCAAAAGCACTTCTTGTGGGGATTTATCTGTGG 2139
Db 7204 TCATTTGACGAGTATCGGCGAGAGACCACTTACTTCTGCGGTGATTCATCTTGAAG 7263
Qy 2140 ACCAGATPACAGGACAGCCCTGACAGATGCGAGACCCCTTAATAAGCTTTTAAAGCTTG 2199
Db 7264 ATTCGTTTACTTCCACAGCGTGTGCGGTGGAGACCTTAAAGAGCTGTTAAAGTTGG 7323
Qy 2200 GAAAACCATTCGCGATGATATCCCAAGATGAGACCGCGCGGCACTGATGATG 2259
Db 7324 GTAAACCGCTCCAGCGGACATAGAGAAAGAGAGAAAGAGAGAGAGAGAGAGAGAG 7383
Qy 2260 AAGCATGCGATGCAAGATTTGAATTTACGAGAGATTTAGTGAAGCGCTGATATCA 2319
Db 7384 AAACAAAGGCGTGTGTTAGATGATATACAGACACTTATGACATGCGCTGCGCACTC 7443
Qy 2320 GATACGATGATATCGGACGCGATCATCACTGCTGCTGCACTTACCGCAAGAG 2379
Db 7444 GGTATGAGGTAGACAAATCACACTGTCTGTGCACTTGAATTTTGTGCGAGACA 7503
Qy 2380 TTAAGAACTTCAGAGCATTAAGAGGAGCCCAATCACCTCTAGCGCTGACCTAAATAG 2439
Db 7504 AAAGGACATTTCAAGCCATCAGAGGGGAAATTAAGCATCTTACGTTGCTTAAATAGT 7563
Qy 2440 TGACGTAGTACA-----CAGCACTTACCACCGGAGAAATTTTCCA 2482
Db 7564 CAGCATATGATCATTTCACTGACTAATPACCAACACACACCATGATATGAGGATTTCT 7623
Qy 2483 TACCTCAGCTGAATTTTCCACAGTTTACCTTCAATCCGATGCTTACCAGATCCA 2542
Db 7624 TTAAATGCTGCGCGCGCTTCCAGGCGCCCATGCTCATGTGAGCGCGGAGAA 7683
Qy 2543 AACCTCTTACAGCCCGCTGAGAGCGCTTGGCGCCCGCTGCTCAATATGAAAT 2602
Db 7684 GAGAGCAGCGCGCCCATATGCTGCGCAAGGACCAACCTTAAGAGAGAGAGAGAGAG 7743
Qy 2603 CTTAAGAGTGTGATGATCACTTCAATTTCAACAAAC--GATCACTTAATCGCGCCAG 2660
Db 7744 CCACAGCGCTGAGTCCCTTAATGATGAGACGCAACATTAACCCCAACGCGCCAC 7803
Qy 2661 GTCCACGCGCAAAAGAAAGAGAGTCTCTTAAGCCAAACCTATCAGCTTAAAGAA 2720
Db 7804 GCCCGCGCGCGCAGAGAGAGGCGCCAAAGCAACCGAAGCGGAAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCCAGAGAGCAAAAC--GCAAGCTTAAACAGGAAAGCAACGTA 2777
Db 7864 AAACACAGAGAAAGAAAGAGAGACCTGCAAAACCCAAACCGGAAAGAGACAGCTTA 7923
Qy 2778 TGTGTATGAAGTTGAGTTCGACAAACATTTCCGATC--ATGCTGAACGCGCAAGTGA 2834
Db 7924 TGGCACTTAATGTTGAGGCGCAGACTGTTCGACGCTCAAAATATGAGGACGAGATGCA 7983
Qy 2835 ATGATATGCTGCTGTGTGTGAGAGAGGCTGATGAAGCACTCCAGCTTGAAGGAAAA 2894
Db 7984 TCGGACACGCACTGCGCATGAAAGAAAGTAATGAACCACTCACTGAAGAACTA 8043
Qy 2895 TTGATTAATGAGCAATTAAGCGCGCTGAATTTGAAGAGGCTAGCATGACATTTGAGT 2954
Db 8044 TTGACACCTCTGTCTATCAAAAGCTCAAAATTTCAACAGTGTGACATATGAGAT 8103
Qy 2955 ACGGCGAGCTTCCCAAGACATGAATCAGACGCTGACATACACGCAACCAACAC 3014
Db 8104 TCGCACAGTTGCGGTCAACATGAGAGAGGCTTCACTTACACAGTGAACACCTTG 8163
Qy 3015 CGGCTTCTAACAATTTGCAACACGCGGCACTGCTGATGAGAAATGAGGAGATTTACCGTAC 3074
Db 8164 AAGGCTTCTAACAATTTGCAACACGCGGCACTGCTGATGAGAGGAGATTTACATTC 8223
Qy 3075 CGAAGAGTGTGGCGGAAAGCGCAACGCGAACAACGATCTTGAACAACAGAGGAGAG 3134

D8 8224 CCCGCGAGTAGAGGAGGAGAGACAGTGGTCGTCGCCGATTATGATTAATCTCAGCGCGG 8283
QY 3135 TTGGGCTATTGTTCTAGAGAGTGCATAATGAGGACCGCTTAACGCCGCTTCACTGTCGA 3194
D8 8284 TTGGCGGATGATCTCGAGGGGCTGATGAGGAAACAAGAACCGCTTTCGGTCTGCA 8343
QY 3195 CTTGGAACCAAGAAAGGGGTGACATTAGGAGTAAACCCCGAAGTTTGAACCGTGGT--- 3251
D8 8344 CTTGGAATAGCAAGAGGAGAGCATCAAGACACCCCGAAGGAGACAGAAAGATGTCG 8403
QY 3252 -----CACTAGTTACAGCGCTATGCGTCTTTCGATGTCAGCTTCCCATGCAAAAC 3305
D8 8404 CTGACCACTGGTGTACCGGCTCATGCTTGCTTGGAAACGTGAGCTTCCCATGCAATCGCC 8463
QY 3306 CACCGGTGCTATTCTCACTGACGCAAGACGACACTGAGTCTCGAAGAGAACGTCG 3365
D8 8464 CGCCCATGCTTACACCGCGAGACATTCAGAGCTCTGACATCTCGAAGAGAACGTCGA 8523
QY 3366 ACAATCCAAATTACGACACGCTGAGAGAACTCTTGAAT---GTCCATCAGCGCGC 3422
D8 8524 ACACGAGGCTTACGACACCTGCTCAACGCTATTTCCGGTGGGATTCGTCGGAGAA 8583
QY 3423 CCAACGAGGATTAACGAGTACTTACACTGACCAAGTCCCTACCTGGGTTTCCCGCT 3482
D8 8584 GTAAAGAGGCTGACCTGACCACTTTTACTTGACAGCCGCTACTTGGGACATGCTCTGT 8643
QY 3483 ATTGCAACACTCAACCGCTGTTTCAAGCCCATTAATAATTGAGAACGTGGGACGAAAT 3542
D8 8644 ACTGTCAACCATATGAAACCGTGTGTTAGCCGATTAAGATGACAGAGCTGAGGATGAG 8703
QY 3543 CTGATGATGATTCGATTAGATTCAGAGCTCTGCGACAAATTCGGCTACAAATCAGGCGCA 3602
D8 8704 CGGACGACACACATACAGATACAGACTTCCGCCAGTTTGAATGACACAAAGCGGAG 8763
QY 3603 CTGCGGATGTCACCAAAATTCGTTACATGCTTTTGACCAAGCAAGCATGCAAGAG 3662
D8 8764 CAGCAAGCTCAAAATTAAGTACCGCTACATGCTGCTGAGCAAGATCATCTCAAGAG 8823
QY 3663 ACAATATGAGAAATATGCTATCAGCAATCTGACACCTGCGCTGCTTGGCCAAAG 3722
D8 8824 GCACCATGAGATGACATCAAGATCAGACCTGAGACCGTGTAGAGGTTATGATCAAAAG 8883
QY 3723 GGTACTTCTGTTGATCAATGTCCTCCAGGTGACAGTGTAAACGCTGATTCACAGCG 3782
D8 8884 GATACCTTCTCTCGCAAGTGTCTCTCCAGGGGACAGGTAAACGTTAGCATACAGGTA 8943
QY 3783 GAGCATCTGAGATTCATGACCGGTGAGAAAGATCAGAGGAAAGTTTGTGGTGAAG 3842
D8 8944 GCAACTGAGCAACGTCAATGACCAATGCGCCGCAAGATTAACCAAAATTCGTGGGACGG 9003
QY 3843 AGGATCTTGTTCACACCGCTCATGGAAGCTGTAAATGCGACGTTTACGATCACT 3902
D8 9004 AAAAATATGACCTTACCTCCGTCAAGGTAAGATTCCTTTCGACATGTACACGCTC 9063
QY 3903 TGAAGAGACGTCGCGGTGATTAACCATGACACAGGCGCACACGCTGTAAAGT 3962
D8 9064 TGAAGAAACACCGCGCGCTACATCACTATGCAAGGCGCGGACCGATTCCTATACAT 9123
QY 3963 CTTATCTGAGAGAAAGCTCAGCGCAAGTGTAAACCACTTTCGCAAGAAAGCTCA 4022
D8 9124 CTTATCTGAGAGATTCATCAAGGAAAGTTTACGGAAGCCACATCCGGAAGAAACATTA 9183
QY 4023 CTTTACGAATGTAAATGTGGCACTACAGACAGATTCGTGAGACCGGAAACGAAGTGA 4082
D8 9184 CGTAGAGTGTAAATGTGGCGCATTAACAAGACGGAACGTTTACGACCCGTACGGAATCA 9243
QY 4083 ACAGCTCACTAAAGCAAGTGCATTTGCTTCAAGAGGACCAAAAGAAATGGGCT 4142
D8 9244 CGGGCTGACCGCCATCAAGAGTGTGCTGCTTATAGAGGACCAAAAGAGTGGGCT 9303
QY 4143 TCAACTGCGCGGATCTTATTAAGGACACAGACCACTCAGTCAAGTAATTTGACATTC 4202
D8 9304 TCAACTGCGCGGATCTGATCAGACACCGGACACAGCGCCAAAGGAAATTTGCAATTC 9363

QY 4203 CATTCCGCTTGACACCGACAGTGTGCCGGTTCGTTAGCTCACAAGCCTTACAGTCAAG 4262
D8 9364 CTTTCAAGCTGATCCGAGTACTGTGATGCTGCTTGTGCCACAGCGCCAACTACTATAC 9423
QY 4263 AGTGGTTCAAAAGGATCAACCTCCACCTGACGTCGATGCGACCAACATTCCTGCAACGA 4322
D8 9424 ACGGCTTAAACATCAGCTCCAAATTAGACACAGACCAATCTGACATTCCTCCACCA 9483
QY 4323 GAAATTTGGGCTCCGAGCAGACGCAACAGCAAAATGGAATTTACAGGTTTACATCCAGA 4382
D8 9484 GGAACATGAGGGCAAAACCCGAAACCAACGAAATGATTCATGGAACACGTTGAA 9543
QY 4383 ATTTTCTGTTGGGCGAAGAGGCTGAGTACGTAATGAGGTTAACATGAACCACTCAGAG 4442
D8 9544 ACTTACCGTGCACCGAGATGCTGTAATATCATATGAGGCAATCAGAACCACTAAGG 9603
QY 4443 TCTGGCCCAAGAGTGGGACACAGGCGACCCCAATGATGAGCGCGCATGATCATATCC 4502
D8 9604 TCTATGCCAAAGATGTGACACAGAGAACCTCAGCGATGGCCACAGAAATAGTACAGC 9663
QY 4503 ACTTATTCATCGGCACTCCAGTCTACCTGTCAATGTGCTGTGTGTGCTGCTTCTTA 4562
D8 9664 ATTAATATCATCGCATCTGTGTACCAATCTTACGCTGATCAGCTGTGTGCGA 9723
QY 4563 TCTGTGAGGCACTGATCATCATCAGAGCTTGATCGCAAGCAAGAAAGAGACTGCTGA 4622
D8 9724 TGATGATTTGGGCTATCTGTTGACAGATTAATGTGCTTAAAGCCGCGGTGAGTCTGA 9783
QY 4623 CGCCATACGCGCTTGACACGCAAGCAAGTACCCACAGATTAAGCGTTTGTGCTGA 4682
D8 9784 CGCCATATGCCCTTGCGCCCAAAATGCCGTGATTCAACTTGCTGCGACTTTTGTGCTGTG 9843
QY 4683 TTGGGCCAACCAAGCTGAAACATTTGGAAGAACTTGAACCATCTGTGTTTAAACAAC 4742
D8 9844 TTAGTGGGTAATGCTGAAACATTCACGAGACCAAGATTTATTATGCTGAAACGCC 9903
QY 4743 AACGTTCTCTGGGACAGTGTGATTCCTCTGAGAGCGCTTGTATTCGTTCCGCT 4802
D8 9904 AGCGTTCTTCTGGGTCAAGTGTATCTCTGCGCGCTGTCTGTCTAATGGCT 9963
QY 4803 GCTTTTCATCTGATGCTCTTTTATTTATGTTGAGGCGCTGTGCTGGGAAAGTACAG 4862
D8 9964 GTTGCTATGCTGTGCTCTTTTATGTTGCTGGCGGCTACTGCGCAAGGTATACG 10023
QY 4863 CTTTGAACATGGAACCATCTGTGCAAAATGTTCCGGGATTCCTGTATTAAGCGTTGCTG 4922
D8 10024 CTTACGAACATGGAACCATCTGTGCAAAATGTTCCAAATGTGCCACAGATACCGTATTAAGCACTGTG 10083
QY 4923 AACGCGAGTTACGCGCACTTAACCTGAGATCAAGTGTCTCATCGGAATTAACAC 4982
D8 10084 AAAGGAGAGGTACGCCCGCTCAATTTGAGATTACTGTCTATGCTCTCGAGGTTTGC 10143
QY 4983 CTTCACTTAAACAGAGTACGTGACCTGCAAAATTCACACAGTATTCCTTACCAACAG 5042
D8 10144 CTTTCAACCAACAAAGATTAATTAATCTGCAAAATTCACCACTGTGCTCCCTCCCTAAAG 10203
QY 5043 TTTAATGCTGCGGTCCCTGAGTGCAGAGCATCTTAAAGGCGGATTAACATGCTCCG 5102
D8 10204 TCAATGCTGCGGTCCCTGGAATGTGACGCCGCTGACGCAATTAATCTGCAAG 10263
QY 5103 TTTTGGCGGTGTATCCCTTTCATGTTGGGAGGCGCAATGCTTCTGACAGTGA 5162
D8 10264 TCTTTGAGGAGGTATACCCCTTCAATGTGGGAGAGACCAATGTTTGTGGACAGTGA 10323
QY 5163 ACACACAATGTAGTGAAGCGTACGTGAGTTCGCTCCAGATGCACTATATGATCACAG 5222
D8 10324 ACAGCAGATGAGTGAAGCGTACGTGAGTTCGCTCCAGATGCACTATGATCACAG 10383
QY 5223 TCGCATTAAGTTTACACAGCTGCTTGAAGTGTGCGCTGCTGTATGATACGCAACA 5282
D8 10384 AGGCGATTAAGTGTACATGCTGCGCGATGAAGATGAGATGCTGTATGTATGACGGAACA 10443

5283 CCACCCGCGACCTGGATACGTTGTCATGCGTCACGCCAGGCTCTCCACGGGACCTGCA 5342
10444 CTACAGGTTTCTTCTAGATGTGTGTAAGCGAGTCAACACGAGGAAAGCTTAAGACCTGA 10503
5343 AGGTATAGGAGGCGGATATAGCGCGCTTTTTCACCTTACCATTAAGTTCGTCATTA 5402
10504 AAGTATAGCTGAGCAATTTAGCATTTGTTTACACATTCATCAAGAGGTCGTTATCA 10563
5403 GAAAGGGGCTTGTATACACTACGACTTCCCTGAGTATGAGCTATGAAACCGAGAGCGT 5462
10564 ATCCGGGCTGTGTATACACTATGACTTTCCGGAATACGAGACGATGAAACCGAGAGCGT 10623
5463 TCGCGATATTTCAAGCATCTCGCTTGTATGATGATACAGACATAGTACCGGCTGACATAC 5522
10624 TTGAGACATTTCAAGCATCTCGCTTGTATGATGATACAGACATAGTACCGGCTGACATAC 10683
5523 GGGTCTGAGGCTTCTGTCAAGAACATCCAGTCCCTTACACCCAGAGATATCAGGCT 5582
10684 GGGTACTAGGCTTCTGTCAAGAACATCCAGTCCCTTACACCCAGAGATATCAGGCT 10743
5583 ATGAATGTGAGAAACAACATCAGAGACGACCCCTGAGAGAAACAGACCATTTGATGTA 5642
10744 TCGAATGTGAGAAACAACATCAGAGACGACCCCTGAGAGAAACGACCATTTGATGTA 10803
5643 AAATTGAAGTGAAGCTCTGCGAGCGCTTACCTGCTTACCGGACATCCCTATCTGCA 5702
10804 AGATTGAGTCAATCCGCTTCAAGCGGTGATGCTGCTATACGGAACATTCCTATTTGTA 10863
5703 TTGATATCCCTGATGACAGCTTTTGTGATGATCATGAAATCACCAATTTTGAAGTTA 5762
10864 TTGATATCCCGAACCTGCTTATCAGACATCAGATGACACATGCTGCTCAGACAGTCA 10923
5763 GCTGACAGTACAGATGATTTATTTGCGAGACTTTGTTGTTGTTCTTCAATTACAT 5822
10924 AATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10983
5823 ACAAAGCTGACAGGAGGAGGACATTTCCAGTTCACTCCCACTCCAGACAGCTGTTTGA 5882
10984 ATGTATCCGACCGGACGAGGACATGCGCTGATCATTCGATTCGACAGCAACCTCC 11043
5883 AGGAAGCCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5942
11044 AAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11103
5943 GCCCAAGCAATTTTATGTTGCTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6002
11104 GCCCAAGGCAATTTTATGTTGCTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 11163
6003 GTAAACCAACCGGACCAATATTTGAGAAACCAATATTTGAGAAATTTCCAGG 6062
11164 GCAACCAACCGGACCAATATTTGAGAAACCAATATTTGAGAAATTTCCAGG 11223
6063 CGGCAATTTCCAAATCTTTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6122
11224 CGGCAATTTCCAAATCTTTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11283
6123 TCATTTGTTGAGGATTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6182
11284 TATTATTTATTTGAGGATTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11343
6183 GACTGAGCGGACCACTGA-CATAGCGGTAATTTGATGATGATGATGATGATGATGATGATGAT 6241
11344 GACCGCTAGCGGACCACTGA-CATAGCGGTAATTTGATGATGATGATGATGATGATGATGATGAT 11403
6242 GCATATGCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11463
11404 GCATATGCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11463
6269 AAAATCGATGATTTTCCGAGAGACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 6318
11464 AAAATCGATGATTTTCCGAGAGACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 11513

RESULT 12
ACU38305
LOCUS 11663 bp RNA linear VRL 30-JAN-1996
DEFINITION Sindbis-like virus isolate S.A.AR86, complete genome.
ACCESSION U38305
VERSION U38305.1 GI:1125069
KEYWORDS
SOURCE
ORGANISM
Sindbis-like virus
Sindbis-like virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; MBV complex.
REFERENCE
1 (sites)
Russell, D.L., Dalrymple, J.M. and Johnston, R.E.
Sindbis virus mutations which coordinately affect glycoprotein
processing, penetration, and virulence in mice
J. Virol. 63 (4), 1619-1623 (1989)
JOURNAL
MEDLINE
PUBMED
2926866
2 (bases 1 to 11663)
Simpson, D.A., Lin, S.C. and Johnston, R.E.
Submitted (11-OCT-1995) Dennis A. Simpson, Microbiology &
Immunology, University of North Carolina, Rm834 FLOB, CB7290,
Chapel Hill, NC 27599-7290, USA
location/Qualifiers
1. 11663
/organism="Sindbis-like virus"
/mol_type="genomic RNA"
/isolate="S.A.AR86"
/db_xref="taxon:4419"
/note="from South Africa."
1. 59
60. 7562
/codon_start=1
/product="nonstructural protein"
/protein_id="AA86135.1"
/db_xref="GI:1125070"
translation="MEKVVNVADVDPGSPFVVOLOKSPFOPEVVAQVTPMDHANARA
FSLASLIELEVPPTTALIDISAPARMSEHOVCMBRSDPMKYASK
AEKAKITNKILHEKIDRLVLDTPDAETSLCEHNDVCTRAEYSWDDVYINAP
GTHIQAKGVRLTYWIGFTDTQFMFSMASSYPAYVNTNMADEKLEKRLGLSTKL
SEGRGKLSIRKELKPSRVSYSVSTGLPEHRAQSOMHLSVFLKQOSYTCR
CDTVSGEYVVKITITSPGIGETVGVAVVNNSEGLKCTVDTVKGERVSPVCTY
IPATIDQMTGIMATDISPDQKLVGLNORIVNGKTNNTNMYLYLPIAOG
SKMKERREDIDNKMTCTREKLTGCGALPRYKXHSFYRPGTORTIVPASFSA
FPMSSVTTSLPMSLRQMKALDPKKEKILQVPEELMEAKAFEDQESRAEKL
REALPLVADGIEBAAEVCEVGLQADTAALVETPRGRVRIIPQANDMIGQYIV
VSPISLVKNALPAHPLADQVKIITHSGRGRVAVEPRDAKVLMPASAVPWEFLA
LSEATLVYNERBEVNRKLYIHANHPAKNTEBOQVTKALELVEVFPVDKRCV
KKEASGVLISGELTNPYHELALGLKTRPAVYKVEITIGVPGSGSAIIRSTV
TABRLNLSGKRENEIEADVLIRGMOITRSKYDVSVMNCHQAVLVYDEARCH
AGALLALAIVRPRKRVLLCGDPKQCGFNNMQLKVNHNPEKDICTRTYKFSRRC
TOPVTALVSTLHYDGKKTNPCKNIEIDTGAETKPKPGDIILTCFGWKOLOIDY
PGEHVTAASQGLTKRGVAVAROVENENPLAITSKVNVLRTBEDRLVTKLOGD
PMIKQLTNPVPGNFQATIEDAEHKGIIAIIINSPARTNPFSCNTKCMAKALEPIL
ATAGIVLTCOMSELPOPADKXHSALYALDVIKIRFEGMDLTSFSGKSIPLTYL
PADBARPRAHNDNSPGRKXYGDVAVAELSRPRVROLAKGTOLDIQTERTVISA
QHNLPVNRNLPHALVPEHKKQGPVEKFSQGRHNSVLVISEKLEAPKRIEWA
PIGIAGADKNKNIAPFPPQARVLDVFNITGKYNHNFQOCEDHAALTKLSRALN
CLNPGGTLVVSXGYADRNSEDVVTALARKVRSVAARPECVSNTEMYLTLFROLDNS
RTROFTPHNLNVLSVYEGTRDVGAPSRYSKRENLADQOEVAVNAAPLGRPE
GVCAIYKRWNSPTDSATENGTAKLTYCOGKQVYHANGPRFRHPRBAELKLOIN
HAYADLVNHNKISVAIPILSTGYAAGKDELVSINCLTLRALDRTADVITCLDK
WKERIDAVLOKESVTEIKDMEIDELVLIHPDSCIKGKSGSTTKGKYSIFEGT
KEHOADAOMEAIKVLFPVDOESNEQLCPLVIGETMEALREKCPDHNSSPPTPLPC
LCMYAMPERRVHRNRNNVKEVTCSSPYLLKIKVQKQCTKVLFNHTPAFLVP
ARKYIEAPEQVPAQAEAPGVAVPTTLPAAINTSLADVDILDMDESSEGLFSS
FSGSDNROVVAVDAVVOEPVPPRLKXMAALAAKROEPTPEASTSADSL
HLSRDGVSISFGSI.PDGMARLAAOPASTCPTDVMHRSFSDGIEIESTRTTES
EPVLPFSRBEPEVNSITSSRAVSFPFRKORRRRSRTREYCLGVGCIYSTGTGPG
HLQKRSVLQNOITPTELRNVLETVAVLDLTSKEQDKLQCMPTPEANSRQSRK
VLENOKAITTERLLSGRLVNSATDQPCYKITYPKPSYSSVPANYPDPRKAVAVCNV
YLHENVPTVASVOYITDEYDALDMDVGTACLDITATFCPAKLSYPRKHEVRAINIS

Query Match	Best Local Similarity	Matches 3751;	Conservative	0;	Mismatches 1985;	Indels	74;	Gaps	9;
580	ACTGACGGTATGAAAGCGGAGCGATATATTTCTCATCGGAAACAGGCCCAAGCTCACTTC	639							
5707	ACTGCTTAAACCGGGGTAGGTGGATACATATTTTTCGACGGAACAAGGCCCTCTGGCACTTGC	5766							
640	AACAGAAATCAGTACGTCATGTAAACTACAGAAGAACTATATTTGGATCGGCGCCCTGCATG	699							
5767	AAAAGAAATCCGTTTCTGCAAGAACACGCTTACAGAAACCACTTGGAGCGCAATGTTCGCG	5826							
700	AGAAATATTAACGCCCGCGCGCTCGATCTCGAAAGAGAGAAATGTTTACAGAAAGAACTTCG	759							
5827	AAAGAAATTAACGCCCGCGCGCTCGACAGTCAAGAAAGAGAAACAGCTCAAACTCAGTAC	5886							

QY	760	AATTATGCGCTCTGGAAGGAAATGAAAGCGAGTATCAATACAGAAAGTGAATAATATGA	819
Db	5887	AGATGATGCTCCACCGAAGCCAAACAAAGCAGATACCACTCTCGAAAGATGAACAAACAGA	5946
QY	820	AAGCAATTACAGCGAGCGAGTCACTTTCTGGATTGGGCACATATCTATTCATCAGAAATGA	879
Db	5947	AAGCATTAACCACTGAGCGACTGCTTTCAAGGCGTACGACTGTATATACCTCTCCACAG---	6003
QY	880	ATCTGTGAGTGTAAAGAGTCAATTATCTGTACCAATCTACTCGTCAACGGTAAATTA	939
Db	6004	ATCAGCCAGATGCTATATAGATCACTACACCGAAACCATCGTATTTCCAGCGTGTACCGAG	6063
QY	940	ACAGTTTACATCTGCAAGAGTCCCGGTAAAGCTGCATTTAGTATTCACAGAAATTT	999
Db	6064	CGAACTACTCTGACCCAAAGTTTGTCTGAGTGTTTGTAAACACTATCTGTGATAGAAAT	6123
QY	1000	ACCCTACAGTAGCCAGTAATATGTATTAACAGATGAATACGATGCTATCTTACATGAGTGG	1059
Db	6124	ACCCGACGGTATGCAATCTTATACAGTACACGACGATAGATGCTTACTTGTGATATGTGAG	6183
QY	1060	ACGGCGCATCTGCTGTCTGTAGATACAGCCACTTTTGTGTGCGGCTTAACTGAAAGTACCC	1119
Db	6184	ACGGGACAGTGCCTTGCTATGATCTGCAACTTTTGGCCCGCCAGGTTTGAAGTTATACC	6243
QY	1120	CAAGGAACATAGCTATTTTGCAGCCAGAGATTAABATCAGCGTCCCATTCGCTATACGA	1179
Db	6244	CGAAAGACACAGATATATAGAGCCCCAAACATCCGACGTGCTTCATACAGCATGACGA	6303
QY	1180	ATACATTTACAAATATGTATATGGCTCAGCTACTATAAAGGAATTTGCAACGTTACCCAAATGC	1239
Db	6304	ACACGTTGCCAAATAGTGTCTACTTTCCCGCATTAATAAGAACTGCACACTCACAAATGC	6363
QY	1240	GAGAAATTAACCTGTCTTATGATTCGGCGGCAATTAATGTTGATTTGTTCAAGAAATACGAT	1299
Db	6364	GTGAATGCCCAACACTGAGACTCAGCGACATTCACAGTTGAATGCTTTGCAAAATATGAT	6423
QY	1300	GCAATGATGATACTGCGGATCTCTTTGGCGATACCCTATTTCCGCTAATCTACAGAAAC	1355
Db	6424	GCAATGACGATATATGGAGAGATTTGCCCGAAAGCCAAATTAAGATCACTACAGATTCG	6483
QY	1360	TTACGCAATATGATACAAAGCTGAAGGGCCGGAAGCGACGACGATGTTTSCGATATCTC	1419
Db	6484	TTACCGCATACGTGTGCCAGACTGAATGAAGCCCTTAAGCCCGCAGCTGTTGCCAAAGAGCG	6543
QY	1420	ATTAATCTAATAACCGTGTGAGAGATTAACAATGATCAATTTGTCATGGATCTTAAAGAG	1479
Db	6544	ATAATTTTGTGCTCCATTTGGTAAGAAAGTGCCTATGATGATGTTGTCATGACATGAAGAG	6603
QY	1480	ATGTCAAAATTAATCTCCCGGACGAAACATACAGAGAGCGGCTTAAGTGTGAGTTATTC	1539
Db	6604	ACGTGAAGATTAACCTGTGACGAAACACACAGAAAGAAAGCCAAAGATACAACTGATAC	6663
QY	1540	AGGCTGACGATCCCTTGTCTACGCGCTTAACCTTTCGCGGATTCATACGCGGAATTTAGCCGTA	1599
Db	6664	AAGCCGCAAGAACCCCTGTGCGACGCTTACCTATCGGGATTCACCGGAGTTAATGTGCCA	6722
QY	1600	GACTGAATGCGGTCCTTCTGCGCAATATACATACTCTCTTTCGACATGTCAAGCGAAGATT	1659
Db	6724	GGCTTACAGCCGTTTGTGATACCCAAACATACAGAGCTCTTTGATGATGTCCGCGGAGGACT	6783
QY	1660	TTGATGCGATTAATTTGCTGAACATTTTCCACACGCGGACCCAGTATTGGAAACGGAATCG	1719
Db	6784	TTGATGCAATCATATAGCAGAAACATTTCAAGCAAGGTGACCCCGTACTGGAGACGGAATTCG	6843
QY	1720	CGTGTGTTGAATAAAGGAGACGACGCTATTCGCAATTTGCGGCTGTGATGATCTGTGAG	1779
Db	6844	CCTGTGTCGACAAAGCCCAAGACGCTATAGGCGTTTACCGGCGCTGTGATCTTTGGAG	6903
QY	1780	ACTTAGGTGTGACCAACCGCTCTTTAATTTGATATGAGGCGCGTTCGGCAATATCAAT	1839
Db	6904	ACCTGGGTGTGACCAACCACTACTGCACTTGATGTAGGTGGCCTTTGGAAATATATCAT	6963
QY	1840	CTGTGCACTTACCTACAGGAAGAGGTTTAATTTGCTGCTCATGATGAATTCGGGATGT	1899

D	6964	CCACCACCTCTGCCACGGGTACCGCTTCAAAATGCGGGCGATGTGAAATCCGGAAGT	7023
Q	1900	TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTGCAGAGTACTACGTG	1955
D	7024	TCCTCACGCTCTTTGTCAACAAGTTTGAATGTGTTATGCCACAGACAGATATTGAGCG	7083
Q	1960	AACGGTTAACACGTCGACGTGCGCGCTCTATGGCGACGATTAACATAGTCATGCTG	2019
D	7084	AGCGGCTTAAACGTCCTCAATGTGACGACTTATTCGGCGACGACAACTTTATACAGGAG	7143
Q	2020	TCGTCCTCGACACTTGTATGGGGGAGAGATGGCGCACTTGGCTGAACATGGAAGTAAAA	2079
D	7144	TAGTATCTGACAAAGAAATGTGAGAGGTGGCCACCTGGCTCAACATGAGGTTAAGA	7203
Q	2080	TTATATGATGCAAGTTATTTGGTATCAAGACACCTTACTCTGTGGGGGATTTTATCTGGATG	2139
D	7204	TCATTGACGCACTCATCGCGCAGAGAACACACTTACTCTTGGGTGGATTCATCTTGGCAG	7263
Q	2140	ACCAGATTAACAGGCAAGCTTGACAGAGTGCAGACCTCTTAAAAAGGCTTTTAAAGCTTG	2199
D	7264	ATTGCGTTTACCTCCACAGCGTGTCCGTGGCGGACCCCTTGAAAAAGGCTGTTAAGTTGG	7322
Q	2200	GAAACCATTTGCCAATTCGATGATACCCAGAGTGGGACCGCGCGCGGCATCTGATATG	2255
D	7324	GTTAAACCGCTCCACACCGACGATGAGCAAGAAAGAAAGAACGACCTCTGCTAGATG	7383
Q	2260	AAGCATCGATGTGAACAGAAATTTGAAATTAACGACGAGTTATGTGAAGGCCGATGAATCA	2319
D	7384	AAACAAAGCGCGTGTATTAGATGATTAACAGACACTTACAGATGCGCTGGGCACTC	7443
Q	2320	GATACGAGATCATCTGCGACGGCTGTATCATCACTCTCTGTCCAAGTTAGCCGAACG	2379
D	7444	GGTATGAGATGAGCAACATCATCACACTGTCCTGCTGGATTTGAGACTTTTGCCAGAC	7503
Q	2380	TTAAGAACTTCAAGAGCATTAAGAGGAGCCCAATACCTCTTACGGCTGACCTTAATGG	2439
D	7504	AAAGGACATTTCAACGCCATCAGAGGGGAAATTAACATCTCTACCGTGTCTTAATAGT	7563
Q	2440	TGACCTAGTAGA-----CAAGCACTTACCACCGGAGATGTTTCA	2482
D	7584	CAGCATATGTAATTTCACTCATCTGATCTAATACACACACACCACTGAATGAGATTTCT	7623
Q	2483	TACCTCAGCTGAATTTTCAACAGTTTACCTTACAAATTCGATGGCTTACCGAGATCA	2542
D	7624	TTAATCATCTCGGCGCGCGCTTCCAGGCCCACTGCATGTGAGGCGCGGAGAA	7683
Q	2543	AACCTCTTAGGCGCGCTGAGAGCGCTTTGGCGCCCGCTGGCTCTCAATTCAGAT	2602
D	7684	GAGGACAGCGCGCCCGATGCTGCGCCCAATGGCTGTCCCAATTCAGCAACTGA	7743
Q	2603	CTTAGAGGTGCAATGTCAACTGTACTTTCAAAACAC--GATCATCTAATCCGCGCCAG	2666
D	7744	CCACAGCCGTCACTGCCCTTCACTGATTTGAGACAGGCACTAGACTTAACCCCAAGCCAC	7803
Q	2661	GTCCACCGCCAAAGAAAGAAAGAGTGTCTTAAGCCAAAACCTTACTAGGCTTAAAAAGA	2720
D	7804	GCCGCGCGCGCGCCAGAAAGAGGCGCCAAACCAACACACAGCCGAAAGAACCA	7863
Q	2721	AGAGACGCAAGCCAAAGAGACGAAC--GCAAGCTTAAACAGGGAAGGACACGTA	2777
D	7864	AAACACAGAGAAAGAAAGAAAGCAACTGCAAAACCAAAACCGGAAAGAGACGCGTA	7923
Q	2778	TGTGTATGAAGTTGGAGTCCGACAAAGACTTTCCGATC--ATGTGAACGGCCAGTGA	2834
D	7924	TGGCACTTAAGTTGAGGCGCACAACACTGTTGACGTCAAAAATGAGGACGAGATGTCA	7983
Q	2835	ATGATATGCTCTGCTGTGCGGAGAGAGCTGATGAACCACTCCAGCTTGAAGGAATA	2894
D	7984	TCGGGCAAGCACTGGCCATGTGAAGGAATGTAAACCACTCACTGAAAGGAATA	8043
Q	2895	TTGATTAATGAGCAATTTAGCGCGCTGAATATGGAAGGCTTAGCATTTGGAGT	2954

Db	8044	TTGACCACTCTGTGCTATCAAAAGCTCAAATTCACCAAGTCTGTACAGATATCAATGAGCT	8103
Qy	2955	ACGGGACGTTTCCCGAAGACATGAATTCAGACAGCTGTCACTATACACAGCACAACCCAC	3014
Db	8104	TGCGACAGTTCCCGGTCAACATGAGAGGTGAGGGGTTTACCTTACACAGGTBAACACCTG	8165
Qy	3015	CGGGCTTTCACACTGGCACAACCGCGCATGTCCAGTATGAGATGGAGATTTACCGTAC	3074
Db	8164	AAGGATTTCAACATGGGACCAACGAGCGGTGCAAGTATGAGGAGGACGATTTACATCC	8223
Qy	3075	CGAGGAGTGGGGGGGAAAGGGCACAAGCCGAAACCGATCTCTGACACAAGAGCCAG	3134
Db	8224	CCCGGGGAGTGGAGGCGAGAGAAACAGTGTCTCTCCATTTATGATTAATCTACAGCCCGG	8283
Qy	3135	TTGTGGCTATTGTTCTAGAGAGGTGCAAAATGAGGCGACGCGTACAGCGCTTTCAGTGTCA	3194
Db	8284	TTGTGCGATAGTCTCTCGAGAGGGCTGATGAGGGAACAAGAACCGCCCTTTCGGTGTCA	8343
Qy	3195	CTTGGAAACGAAAAGGGGTGACATTAGAGATACCCCGAAGGTTCTGAAACCGTGT--	3251
Db	8344	CCTGGAAATAGCAAAAGGGAAGACAATTAAGACAACCCCGAAGGGAAGAAAGATGTCTG	8403
Qy	3252	-----CACTAGTTACACCGCTATAGCGGTCTTGGAAATGTACGTTCCCATGCGACAAC	3305
Db	8404	CTGCACCACTGTGTACCGCCATGTGTCTTGTGAAACGTGAGGTTCCCATGTACATCGCC	8463
Qy	3306	CACCCGTGTGTATTACCTGACGCGCAGAACCAACACTGACGTCGTGCAAGAGAAGTGC	3365
Db	8464	CGCCCACTGTGTACCCCGGCAACCACTCAGAGCTCTGTGACATCTCTGAGAGAAAGTGA	8523
Qy	3366	ACAATCCAAATTAAGACACGCTGTGTGAGAACGTTTGAAAT--GTCATACACGCGGC	3422
Db	8524	ACCAGAGGCTTACGACACCTGTCTCAACGCGCATATTGCGGTGGGATTCGCGGAGAA	8583
Qy	3423	CCAAACGAGAATTACGATATCTTGCACGTACCACTCCCTACTGGGGTTCTGCCCT	3482
Db	8584	GTAAGAAAGAGGTACATGACGACCTTTACCTTGAACAGCCCGTACTTGGGCAATGCTGT	8643
Qy	3483	ATTGACAGACCTCAACGCGCGTGTTCAGCCCAATTAATAATTGAGAACGTGTGGAGCGAAT	3542
Db	8644	ACTGTACCACTACGAAACCGTGTCTTGAAGCCGATTTAGATGAGACAGGTCTTGGGATGAG	8703
Qy	3543	CTGATGATGATGATTAAGAAATCCAGGTCTTCGACCAATTCGCGTACAACTCAGCAGGCA	3602
Db	8704	CGGACGACAAACCAATACGATACAGACTTCGCGCCAGTTTGATTAAGACCAAGCCGAG	8763
Qy	3603	CTGCGGATGTACCAAAATTCGTTACATGTCTTTGACACAGACCACTGACATCAAGAG	3662
Db	8764	CAGCAAGCTCAAAATAGTACCGCTACATGTGTGCTCGAGCAGGATCAATCTGTCAAGAG	8823
Qy	3663	ACAGTATGGAATAATAGCTATGAGCATGTGAGACCTGCGGTGCTTGGCCACAAG	3722
Db	8824	GCACCATGATGACATCAAGATGACACTTAGAGACGTTGTAAGGCTTAGCTTACAAAG	8883
Qy	3723	GGTACTTCTGTGTGCTCAATGTCTCTCAGGTGACAGTGTAAACCGTCAAGTATCAAGAGC	3782
Db	8884	GATACTTCTCTCTGCGAAGTGTCTCTCAAGGAGCAAGGTAAACGTTAGCATACGAGTA	8943
Qy	3783	GAGCATCTGAGAATTATGACACCGTGAAGAAAAGATCAGAGGAAGTTGTTCGTTAG	3842
Db	8944	GCAACTCAGCAACGTCAATGCAATGCGCCGCAAGATTAACCAAAATTCGTGTGGAAGGG	9003
Qy	3843	AGGAGTACTTGTCCCAACCGGTGATGAGAAAGTGTAAAGTGCACGTTTACGATCACT	3902
Db	9004	AAAAATATGACTTACCTCCGTTACGTTAAGAAAGATTCCTTGCACAGTGTACACCGTC	9063
Qy	3903	TGAAGAGACGTCGCGGGGTATCAATCAATGACAGGCGACGCCACACGCTATAGT	3962
Db	9064	TGAAGAAACAAACCGCGGGCTATCATCTATGACAAAGCGGGACGCCATGTCTATCAT	9123
Qy	3963	CTTATCTGAGGAAGCGTCAAGCGAAGTGTACATTAACAACCTTTCTGCGACAGACGTCA	4022
Db	9124	CTTATCTGAGGAATCATCAGGGAAGTTTACGGAAGCCACCATTCGGGGAAGAACCTTA	9183

OY	4023	CTTACGATGTTAAAGTGTGGCATCTACAGCACAGAGTTATCTGTAGACACGGACAAGAAAGATGA	4082
Db	9184	CGTACGAGTCCAAAGTGGGGGGAATTAACAAAGACCGGAAACCGTTACGACCCGCTACGAAATCA	9243
OY	4083	ACGAGCTGACATTAAGCAAAACAGTGCATTCTCTACAAAGACGCCAAAGAAATGGAGTCT	4142
Db	9244	CGGGCTGCACCGGCATCAGACAGTGGCTGGCTTTAAGACCAACAAAGAAATGGAGTCT	9303
OY	4143	TCAACTGCGCGGATCTTATTAGGCACACAGACCACTCAGTGCAGAGTTAAATTCACATTTC	4202
Db	9304	TCAACTGCGCGGATCTGATCAGACACGCCGACACACACGCCCAAGGAAATTCACATTTC	9363
OY	4203	CATTCCGCTTGACACCGACAGTCTGGCCCGGTTCCGTTAGCTACACAGCCTACAGCTACAGCA	4262
Db	9364	CTTTCAAGACTGATCCGAGATACCTTCGATGGTCTCTGTGTCACACGGCGGAAAGCTGATC	9423
OY	4263	AGTGGTTCAAAAGGCATCACCTTCCACTGCACTGCAATGCGACCAATTCGTGCAACAGCA	4322
Db	9424	ACGGCTTTAAACATCATCAGCCTTCCAATTAGACACAGCCATCTGACATTGCTCACACACA	9483
OY	4323	GAAATTTGGGGCTGCCAGCAGACGCCAACACAGATGAGATTACAGGCTTCAATCCAGCA	4382
Db	9484	GGACACTAGGGGGCAAAACCCGGAAACCAACCTAAGATGATCAACGGAAACAGGTTAGAA	9543
OY	4383	ATTTTTCGTGGGGCGAAGAGGGGTGGAGTACGTATGGGGTAAACAATGAACAGTCAAG	4442
Db	9544	ACTTCACCTTCGACCCGAGATGGCTGTGAATACATATGGGGCAATACAGAACCACTGAAGG	9603
OY	4443	TCCTGGGCCCAAGAGTGGGCAACAAGCGCAACCAATGATGTCGGCCGACATGAGATCATTC	4502
Db	9604	TCCTATGCCCAAGATCTGCACACAGAGAGCCCTCACGATGGCCACACGAAATAGTACAGC	9663
OY	4503	ACTATTTTCATCGGCATCCAGTCCATCACTGTCAATGTGATGTGTCGTGGCTTGTGCTA	4562
Db	9664	ATTACTATCATCGCCATCTGTGTACCACTTTAGCCGTGCGATCAGCTGCTGGGCA	9723
OY	4563	TCCTGTGAGGCATGTCATCAGCAGCTTGATTCGCCAAAGCAAGAAAGACTGCTCTGA	4622
Db	9724	TGATGATTTGGGCTAACTGTGCAGCACTTATGTCTGTAAAGCGCCGCTGAGTGGCTGA	9783
OY	4623	CGGCATACGGCCTTGACACCGAACCGCAAGGTTACCAACAGATTAGGGGTTTTGTGTGA	4682
Db	9784	CGGCATATGCCCCGTGGCCCCAAATGCGGTGATTCCACTTCGCTGGGACTTTGTGTGTGG	9843
OY	4683	TTCCGGCCAAACGCTGTGAACATTTGGAGAAACCTTTGAAACCACTGTGTGTTTAAACACC	4742
Db	9844	TTAGGTGGGGCTAAAGCTGAAACATTCACCGAAGACATGATTACTTAATGTTGAACAGCC	9903
OY	4743	AAACGTTTCTGTGGGCACAGTTGTGCAATTCCTGTGGACGGCTTGTATTCTGTTCGGCT	4802
Db	9904	AGCGGTTCTTGTGGGTCAACCTGTGTATACCTGTGGCCGCTGTGTGTGTCTTAATGCGCT	9963
OY	4803	GCTTTTCAATGCTGATGCTTTTATTTTGTGTGAGGCGCTGCTCGTGGGAAAGGTGAGAC	4862
Db	9964	GTTGCTCATGCTGCTGCTGCTTTTATTTAGTGTGTCGGCGCTTACTCTGGGAAAGGTGAGAC	10023
OY	4863	CCTTTCGAACATGCACCACTGTGCAAAATGTTCCGGGGAATCCGCTAATGAAGCGTGTGTCG	4922
Db	10024	CCTACGAACAATGCACCACTGTTTCCAATATGTGCCACAGATACCGTATTAAGGCACTTGTTCG	10083
OY	4923	AAACGGCAGGTTAAGCGGCCACTTAACCTGTGAATACACGGTGTCTCAATGTGGAAATTAACAC	4982
Db	10084	AAAGGGCAGGGTACGCCCCGCTCAATTTGGAGATATGTGCAATGCTCTCGGAGGTTTTC	10143
OY	4983	CTTCAACTAACAAAGAGTACGTGACCTGTCAAAATTTCCACACATCATCTCTTCCACCAAG	5042
Db	10144	CTTCCACCAACCAAGAGTACATTAACCTGCAAAATTCACACATGTGTGTCCCTCCCTTAAG	10203
OY	5043	TTAAATGCTGGGGTCCCTCGAGTGCAAAGCATCTCTAAAGCGGATTAACATGTCGCG	5102
Db	10204	TCAATATGTGGCTCTCTTGGAAATGTAGCCCGCGCTCACACGACATTAACCTTCGAAG	10263

QY	5103	TTTTTGCGGCTGTGTACCTCTTCACTGTGGGGAGGCGCACATGCTTCTGTGACAGTGTAGA	5162
Db	10264	TCCTTGTGGAGGGGTGTACCCCTTCTCACTGTGGGAGAGACACATGTTTTTGTGGACAGTGTAGA	10322
QY	5163	AACAACAACCTAGTAGGAGGCGGTACGTGTCAAGTTTCGCTCAGACTGTACATTTAGATCAGCGAG	5222
Db	10324	ACAGCGACATAGTAGAGCGGTACCTCAATTTGTCTAGTAGATTGGCGACCTGTACCACGCGC	10308
QY	5223	TCGCACCTAAAGTTTACACAGACTGCTGTGAAGTCGGCGCTGTGTATGTATGTATACGGCAACA	5282
Db	10384	AGGGGATTAAAGTGTCATCTGTGCGCGCGCATGAAGTAGAGACTGTCCGTTATGTGTACGGGAACA	10404
QY	5283	CCACCGCGGCACCTGTGGATACGTTTGTCACTGCGCTCACCGCAGGTTCTCACGGGACCTGTGA	5342
Db	10444	CTACCAAGTTTCTCTAGATGTGTACCTGTGAACGGAGTCACACACAGGAACCTCTTAAAGACCTGTGA	10500
QY	5343	AGGTCATGTGACGGGCGCATATACGCGCTTTTACCCCTTTGACCATTAAGGTGTCTATCA	5402
Db	10504	AAGTCATAGCTGTGACCAATTTACACATTTGTTTACACATTTGATCACAAGGTGCTTTATCA	10566
QY	5403	GAAAGGGGCTTGTTTACAACTACACACTTCCCTGTAGTAGTGAGCTATGAAACAGAGACGT	5462
Db	10564	ATCCGGCGCTGTGTACACTATGACTTTCCGGAATGTGGAGGCGATGAACACAGAGACGT	10622
QY	5463	TCGGGATATTTCAAGCATCTCTCGCTTATGTCTACAGACATGTATGCCGCACTGTACATAC	5522
Db	10624	TTGGAGACATTTCAAGCTTACCTCTTGTGACTGTGACAAAGACCTCATGTCCAGCACAGACATTTA	10682
QY	5523	GGCTGTGTGAACCTCTTGTGTGAAGACATCCACGTCCTTACACCCAGCAGTATCAGAGT	5582
Db	10684	GGCTACTTAACCTTCCGCAAGAACGTGTATGTCCGTGTACGCAAGCGCGCATCTGGAT	10742
QY	5583	ATGAAATGTGAAGAACCACTCAGACGACCCCTGTCAAGAAACAGACCAATTTGGATGTGA	5642
Db	10742	TCGAGATGTGAAGAAACAACTCAGCGCGCCACACTGTGACAGAAACCGCCCTTTTGGGTGCA	10802
QY	5643	AAATTTGAAGTAGAGCCTCTGTGAGCGTCTTAACTGTGCTTACGGGACATCCCTATCTGTGA	5702
Db	10804	AGATGTGACGTCAATCCGCTTGTGACCGGTGTGACTGTCTATACGGGAACATTTCCCTATTTCTTA	10862
QY	5703	TTGACATCCCTGTATGTGACGCTTTTGTGATGATCATCAGAAATCACCAACAATTTTGAAGTTTA	5762
Db	10864	TTGACATCCCAACGCTGTCTTTATTCAGGAATGATGATGACCACTGTGTCTCAACACTCA	10922
QY	5763	GCTGCACAGTAGACAGACTGTCAATTTATTTCTGTGACACTTTGTGTGTTCTCTAACTTAAGT	5822
Db	10924	AATGTGATGTGATGTAGTAGTGACTTATTTACGGGACTTCGGAGGGAATGGCTTACCTGTGAGT	10982
QY	5823	ACAAAGCTGACAGGAGGAGGACATTGTCTCAAGTTCACTCCACTCCACGACAGCTGTTTTGA	5882
Db	10984	ATGTATCCGACCGGAGAGGACAAATGCCCTGTATCTTCGACTTTCGAGCAGACAACTCTCC	11042
QY	5883	AGGAAGGCAACCACTATGTGACTGTCCGTAGGACAGATTAACATCTATTTTATAGCANTGTGA	5942
Db	11044	AAGAGTCCACAGTTCACTGTCTGTGAAGAAAGAGCGGTGACAGTACACTTTCAGCACCGCAGA	11102
QY	5943	GCCCAACAGCAAAATTTTATAGTTTTCGCTATGCGGCAAGAAAGTCCACCTGTCAATSTCAAT	6002
Db	11104	GCCCAACAGGCAACTTCAATTTATGTGTGTGTGTGTGAAGAAACAACTGTCAATGTCAAGAT	11162
QY	6003	GTAACCAACCGGCGGACCACTAAATTTGGAAGAACCACTAAAGTGTGACCAAGATTTCCAGG	6062
Db	11164	GCAAAACCAACGCTGTATCATATGTGTGACACCCGCGCAAAAAAATGACCAAGAAATTTCCAGG	11222
QY	6063	CGGCAAGTTTCCAAAACATCTTGGAACTGTGCTGTGTGACTGTGTTTGGGGAGCATCATCC	6122
Db	11224	CCGCAATCTCAAAAACTTCACTGAGGATGTGGCTGTGTGCTTTTTCGGGCGCGCTCTGTGCGC	11282
QY	6123	TCAATGTGTGTGAGCTTATAGTGTGTGTGTGCTGTGACTGTATAGCTTAAACAACAGTGTAGT	6182
Db	11284	TATTAATATTATGTGACTTATGTATTTTCTGTTCGACATGATGTGTGACTGTGACAGAAAGAT	11342
QY	6183	GACTGTAGCGCGCAACTGTGA-CATAGCGGTAAATATCATGTATCTTCCAGGAGGCTGTGT	6242

Db 11344 GACCGCTACGCCCAATGACCCGACAGCAAAATCGATGTAATCCGAGGAATGATGT 11403
Qy 6242 GCATATGCGACGCGCGCT-----TGACACT 6268
Db 11404 GCATATGCGATAGCGTGTATATTAGATCCCGCTTACCGCGGCAATATAGCAACAC 11463
Qy 6269 AAAACTGATGATTTTCGAGGAGACAGTGCATATATGCTGTGAGTGT 6318
Db 11464 AAAACTGACGATTTCCGAGGAGCGAGTGCATATATGCTGTGAGTGT 11513

RESULT 13
AR042412 11703 bp DNA linear PAT 29-SEP-1999
LOCUS AR042412
DEFINITION Sequence 8 from patent US 5811407.
ACCESSION AR042412
VERSION AR042412.1 GI:5962908
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11703)
AUTHORS Johnston,R.E., Davis,N.L. and Simpson,D.A.
TITLE System for the in vivo delivery and expression of heterologous genes in the bone marrow
JOURNAL Patent: US 5811407-A 8 22-SEP-1998;
FEATURES
SOURCE location/Qualifiers
BASE COUNT 3306 a 3051 c 2913 g 2433 t
ORIGIN

Query Match 36.8%; Score 2388.4; DB 6; Length 11703;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3698; Conservative 0; Mismatches 1981; Indels 41; Gaps 8;

Qy 580 ACTGACGGTATGAAAGCGGAGCGTATATTTCTCATCGGAAACAGGCCAGGTCACTTC 639
Db 5746 ACTGACTAACGGGGTAGGTGGTATATTTTCACGACACAGGCCCTGGGCACTTGC 5805
Qy 640 AACAGAAATCGAGTCAATGTAATTAACAAAGAACTATATTTGATGGGCGGTCCATG 699
Db 5806 AAAAAGATCCGTTCTGCAAGAACGACTTACAGAACCGACTTGGAGCGCAATGTCCTGG 5865
Qy 700 AGAATATTAGCCCGCGCCCTCGATCTCGAAAGAGAAATGTTACGAAGAACTGC 759
Db 5866 AAAAGATTATGCCCCCGGTCTGACAGTCGAAAGAGAGAACACTCAAACTCAGGTACC 5925
Qy 760 AATTATGCGCTCTGTAAGAAATAGAGAGAGGTATCAATCAGAAAGTAGAAATATGA 819
Db 5926 AGATATATGCCACCGAAGCAAAAGTAGTACAGTCTGTAAGTAGAAATTCAGA 5985
Qy 820 AAGCAATTACAGCGGAGGAGCTATTTCTGATTGGGCAATATCTATCATCAAGTGA 879
Db 5986 AAGCATATACCACTAGCGACTACTGTGAGACTGACACTATTAATCTGACAG--- 6042
Qy 880 ATCTGTGAGTGTATACAGAGTCAATTTCTGTACCAATCTACTCTGTCACGGTAATTA 939
Db 6043 ATCAGCCAGAAATGCTATAGATCACCTATCCGAAACCAATTGTATCCAGTACCGTACC 6102
Qy 940 ACAGGTTTACATCTGACAGAGTGCAGGTTTAAACGTGCACTTATTTTCAAGAGATT 999
Db 6103 CGAATCTACTCCGATCCACAGTTCCGTGTAAGTCTGTATACAACTATCTGACGTGAAT 6162
Qy 1000 ACCCTACAGTAGCCAGTTATTTATTAACAGATGAATAGATCGTATCTTGACATGTGG 1059
Db 6163 ATCCGACAGTAGCATCTTATGATTAATGACGAGATGATGATCTTAATGATGATG 6222
Qy 1060 ACGGCGCATCGTGTCTGATATACAGCACTTTTGTCCGCTTAACTGAGAGCTACC 1119
Db 6223 ACGGACAGTGCCTGCTGATGACTGCAACCTTCTGCCCCGCTAAAGCTTGAAGTATAC 6282

Qy 1120 CAAGAGACATAGCTATTTGACGCCAGAGATAGATCAGCCGCTCCATCGCTATACAGA 1179
Db 6283 CGAAAAAATAGAGTATAGAGCCCGAATATCCGAGTGCCTTCATCAGGATGACAGA 6342
Qy 1180 ATACATTCAAAATATATATGCTGACGACTACTTAAAGAAATTTGCAAGCTTACCCAAATGC 1239
Db 6343 ACAGCTTCAAAATATGCTGATTTGCGCACTTAAAGAAATTTGCAAGCTTACCGAGATGC 6402
Qy 1240 GAGAAATTCCTGTCTAGATTTCGCGGCGATTTAATGTTGATTTGTTCAAGAAATACGAT 1299
Db 6403 GTGAATCTCCAACTGAGTCTCAGGACATTCATATGTGAATGCTTTGAAAATATGAT 6462
Qy 1300 GCAATGATGATGATGAGTACCTTTGCGGATTAACCTATTCCGCTTAACTACAGAAAG 1359
Db 6463 GTATATGAGATGATGAGAGAGTTCGCTCGGAGGCCAATTAAGTTCACATGAGTTTG 6522
Qy 1360 TTACGCAATATGTACAAAGCTGAAAGGGCCGAAAGCAGCAGATTTGTTGCAATATCTC 1419
Db 6523 TCACCGCATATGTAGCTGAGTGAAGGCCCTTAAGCCGCGCACTATTTTCAAGAGACT 6582
Qy 1420 ATTAATCTAAAACGTTGACGAGATACCAATGATCAATTTGCTCATGATCTAAGAGAG 1479
Db 6583 ATTAATTTGCTCCATTGCGAAGAAAGTCTTATGATGATGCTGTCATGACATGAAGAG 6642
Qy 1480 ATGTCAAAATTAATCCCGGACAGAAACATACAGAGAGCGGCTTAAAGGTGAGGTTATTC 1539
Db 6643 ACGTAAAGATTACACAGCAGCAGAAACACAGAAAGAAAGCCGAATATACAGTATAC 6702
Qy 1540 AAGCTGACGATCCCTTGTCTACCGTTACTTTGCGGAGTCCATCGGGAATTAATGCTGTA 1599
Db 6703 AAGCGCAGAAACCCCTGCGACTCTTACTATGCGGAGTTCACCGGGAATTAAGCCGTA 6762
Qy 1600 GACTGAATGCGGCGTCTGCAAAATATCCATCTCTTGGACATGTCACGCGAAGATT 1659
Db 6763 GACTTACGCGGCTGCTGTTCCAAACATTCACAGCTTTTGAATGTCGCGGAGATT 6822
Qy 1660 TTGATGCGATTTATGCTGAAATTTTCACCAACGCGGACCGAGTATGGAACCGACATCG 1719
Db 6823 TTGATGCAATATAGCAAGAACACTTCAAGCAAGGAGACCGGTAATGAGACGATAC 6882
Qy 1720 CGTGTGATTAAGAGCAAGACGAGTATGCGCAATTTGCGGCTGATGATCTTGAAG 1779
Db 6883 CATCATTCGACAAAGCCAAAGCAGCAGCTATGCGTTAACCGGCTGATGATCTTGAAG 6942
Qy 1780 ACTTAGGTGACCAACCGCTCTTAATTTGATGAGAGCGCGCTTGGCAATATCAAT 1839
Db 6943 ACTGGGTGTGATACACCTACTGACTGATGATGAGTGGCGCTTGGAGAAATATAT 7002
Qy 1840 CTGTGACCTTACTTACAGGAACGAGTTTAAATTTGTCGATGATGAATCCGGTATGT 1899
Db 7003 CCAACCATCTACCTACGGGTACTGCTTTAAATTCGGGGCATATGAAATCCGAATGT 7062
Qy 1900 TCTTAACGCTGTGTTCAACACACTAGTCAATATCATGATTTGCTAGCAGAGTACTACGTG 1959
Db 7063 TCTTCACACTTTTGTCAACACAGTTTGAATGCTATGTCGACGAGATACAGAAG 7122
Qy 1960 AAGGTTTACACGCTCAGCGTGCAGGCTCTATTCGGGAGATTAACATAGTGTGTG 2019
Db 7123 AGCGCTTAAACGTTACAGATGTCAAGGTTTATGCGGAGACAACTCATATGAG 7182
Qy 2020 TCGTCTCCGACACTTGTAGCGGAGATGCGCACTTGGCTGAAACATGGAATTAATA 2079
Db 7183 TAGATCTGACAAAGAAATGCTGAGAGTGCACCTGCTCAACATGAGGTTTAAGA 7242
Qy 2080 TTATGATGAGATATGATGTTCAAGACCTTACTTGTGCGGGAATTTATCTGTGTG 2139
Db 7243 TCATGACGAGTACGTAGGTGAGAGACCACTTACTTGTGCGGGAATTTATCTTGAAG 7302
Qy 2140 ACCGATTAACAGGACAGCTGCAAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7303 ATTGGTTACTTCAAGACCGTGCAGCGGAGGAGCCCTGTAAGAGCTGTTTAAAGTTGG 7362
Qy 2200 GAAACCATTCGAGTGAATACCCAAAGCTCGAACCGCGCGGCACTGATGATG 2259

Db	7363	GTAAACCGCTCCACGCCGACGACGACGAAGACGAAGAGACGCGCTCTGCTAGTGG	74222
Qy	2260	AAGCAATGCGATGGAACAGAAATTTGGAAATTAACGACGAGTTGATGAAAGCCGTAGAAATCA	2319
Db	7423	AAACAAAGGCGGTGGTTTAGAGTAAGGTAAACAGGACACTTTAGCAGTGGCCGTGAAGACACC	7482
Qy	2320	GATACGAGATCATACTGGCAGGCGCTGATCATCA GGTCTGTCTCCAGCTTACCGAAAGCG	2379
Db	7483	GGTATGAGGTAGACAATTTTACCTCTGCTCACTGGCAATTGAAACTTTTCCAGAGACA	7542
Qy	2380	TTAAGAACTTCAAGAGCAATAAGAGGAGACCAATACCCCTTAAGGGCTGACCTTAATAG	2439
Db	7543	AAAGGCAATTCGAAGCCATCAGAGGGGAAATTAAGCAATCTCTACGGTGGTCTTAAATAGT	7602
Qy	2440	TGACGTAGTACACGACACTTAC-----CCACGGCAGAAATGTTTCCATACC	2486
Db	7603	CAGCATAGTAACATTTCACTCTGACTTAATCTAACACACACACACATGAAATAGAGATTCT	7652
Qy	2487	CTCAGCTGAATCTTCCACCAAGTTTACCTTAACAAATCCGATGGCTTACCAGATCCAAAC	2546
Db	7663	TTAATCATGCTCGGCGCGCGCCCTTTCCGGGCCCCCACTGGCATGTGAGGCGCGGAGAA	7722
Qy	2547	CTCCTAGCGCCGCTGGAGGCGCTTTGGGCCCCCGCTGGCTGCTCAATTCGAAGATCTTA	2606
Db	7723	GGAGGCAAGCGGCGCCCGATGCGCTCGCCGCAAGGCGCTGGCTTTCAAATTCAGACACTGA	7782
Qy	2607	GGAGGTGATATGTCACATTTGACTTTCAAAACAACGATTCACATATCCGCGCGCAGGTCCA	2665
Db	7783	CCACAGCCGTCAAGTCCCTTACTACTTGAACAGGCAACTAGACTCAACCCCAAGTCCAC	7842
Qy	2666	-----CCGCCAAGAGAAAGAAAGTGTCTCTAAGCCAAACCTTACGCTTAAAGAA	2720
Db	7843	GCCGCGCCACCGCGCCAGAAAGACGAGCGCCCAAGCAACACCGAAGCCGAAGAAACCA	7902
Qy	2721	AGAAAGCAAGCCCAAGAGACGAAAC--GCAAGCTTAACTCAGGGAAAGCAACGTA	2777
Db	7903	AAAGCGAGAAAGAAAGAAAGCAACTGTGAAACCCAAACCCGGAAAGGACACGCGCA	7962
Qy	2778	TGTGTATGAATTTGGAGTCCGACCAAGCAACTTTCGATCATG--CTGAACGGCCAAAGTGA	2834
Db	7963	TGCACTTTAATTTGAGGCGCCAGACTTTTTCGACGTACAGTAACAGGACCGAGAGTGTCA	8022
Qy	2835	ATGATATATGCTCGGTGTCTCGAGAGAGGCTGATGAAACCACTCCACCTTTAAGAGAAATA	2894
Db	8023	TCGGGCAAGCACTGGCCCATGGAAGAAAGTAATGAATACCTGTGACGTGAAGAGAACCA	8082
Qy	2895	TTGATATGACCAATTAGCGGCGGTGAATTGAAGAGGCTAGCAATGTACATTTGAGCT	2954
Db	8083	TCGACACACCTGTGTCTATCAAAAGCTCAAAATTTTACCACTGTCTCAGCAATACACTGAGT	8142
Qy	2955	ACGGGACGTTCCCGAGAACATGAAATTCAGACAGGCTCAGTAACCCAGGACAAACAC	3014
Db	8143	TGCACAATTTCCCACTCAACTGAAATGAGGACCTTACCTTACACCGTAGAACACCCCG	8202
Qy	3015	CGGGCTTCTCAACTGGCACCACCGCGCAGTCCAGTATGAGAAATGGAGATTTTACCGTAC	3074
Db	8203	AAGGATTTCTAATACGTGGCACACCGGAGCGGTGCAGTAAATGAGAGTATTTACCAATC	8262
Qy	3075	CGAGAGAGTGGGCGGGAAGGCGCACGCGGAAGACCGATCTTGGACAACAGAGGCAAG	3134
Db	8263	CTCGGGAGTGGAGGCGAGAGAGACACGGTCTGTCGATCATGTATTAATCTCCGTCGGG	8322
Qy	3135	TTTGTGCTATTTGTTCTAGAGGTGCAATGAGGAGGAGACGCGTATTCAGTGTGCA	3194
Db	8323	TTTGTGCGATATGCTCTCGTGTGAGCTGATGAAAGAACAGCAATGTCCCTTTGTGGTGTCA	8382
Qy	3195	CTTGGAAACCAAAAGGGGTGACACTTAGAGATACCCCGAAGGTTCTGAACCGTGT---	3251
Db	8383	CCTGGAAATGTAAGGGAAGCAACTTAAGACAGACCCCGGAAGGGAAGAAAGTGTGTCCG	8442
Qy	3252	-----CACTAGTTACAGCGCTATGCGTCTTTCGATGTCAAGTTCCTCATGCGACAAAC	3305

Db	8443	CAGCACCACTGCTACGGGCAATGTGTTGCTCGGAAATGTGACCTTCCCATGCGACCGCC	8502
Oy	3306	CACCCTGTGTCTATTCTACCTGACCGCAGAAAGCAACCTCTGCAGAGTCTCGAAGAGAACTGTG	3365
Db	8503	CGCCCACTAGTGTATACCGCGGAACTTCCAGAGCGCTCGACATCTCTTGAAGAGAACTGTA	8562
Oy	3366	ACAAATCCAAATYACGACACGCTGTGGAGAACGTTGMAATGTCCATCAC--GCGGC	3422
Db	8563	ACCAATGAGGCTTACGATACCTCTGCTCAATGCAATATGTCGGGTGAGATCGTCTGACAGAA	8622
Oy	3423	CCAAACGAGCACTTACGATGATCTTCACTGACAGTCCCTCACTCGGGGTTCTGCGCGT	3482
Db	8623	GCAAAAGAGGCTACCTGACGACTTTTACCTGACACAGCCCTACTTTGGGACATGCTCGT	8682
Oy	3483	ATTGCAAGACACTAACCGCGCTGTTTACGCCCAATTAATAATYAGAACGCTGTGGACGAAT	3542
Db	8683	ACTGCGACCAATCTAGAACCGTGTCTTACGCCCTGTTAATGATGACAGAGTCTGGACGAAG	8742
Oy	3543	CTGATGATGATCGATTGAATTCAGGTTCTGGGACAAATTCGCTACAAATYAGGCAGGCA	3602
Db	8743	CGGACGATACCAACCATAGCATACAGACTTCCGCCAGTTTGATATGACACCAAGCCGAG	8802
Oy	3603	CTGGGAAATGTACCAAAATTCGGTTACATGTCTTGCACACAGACCATGACATCAAGAG	3662
Db	8803	CAGACAGCGCAACCAAGTACCGCTCAATGTCTTGGACAGATACACCTTTAAAGAG	8862
Oy	3663	ACAAATATGAGAAATYATGCTATACGACATCTGAGACCTCGCGCTGTCTTGGCCACAAG	3722
Db	8863	GCACCAATGAAATGATCAATCAAAATYATGACACTGAGCACCGTGTAGAAAGGTTTAGTACAAG	8922
Oy	3723	GGTACTTCTGTGCTCATATGTCTTCCAGGTGACAGTGTAAACCGTCAATATCACAGCG	3782
Db	8923	GATCTTTCTCTCGCAAAATGCCCCCTCAGGGGACACGTTAACGTTAGCATGTGAGTA	8982
Oy	3783	GAGCATCTGGAATTCATGACCGGTGAGAAAAGATCAGAGGAAGTTTCTCGGTAG	3842
Db	8983	GCAACTCAGCAACGCTCATGTACTGTGCGCGCAAGATTAACCAAAATTCGTGGAGCGG	9042
Oy	3843	AGGAGTACTTGTCTCCACCGCTCCATGSAAGCTGTAAAGTGCACGTTTACGATCACT	3902
Db	9043	AAAAATATGATCTTACCTTCCGTTACCGGTAAABAAATTCCTTGACAGTGTACAGACGTC	9102
Oy	3903	TGAAGAGACGCTGTGCGGGGTACATTAACATGACACAGGCCAGGCCACACGCGTATTAAT	3962
Db	9103	TGAAAGAAACACTGCAAGGCTTACATCACTATGACACAGCGCGGACCGCACCTTATATAT	9162
Oy	3963	CCATATCTGAGGAAGCGTCAGGGCGCAAGTGTACATTAACAACCTTCTGGGACGAAGCGCA	4022
Db	9163	CTTACTCTGGAAGATCATCAGGGAAAGTTTACGAAAGCGCCCATCTGGGAGGAACATTA	9222
Oy	4023	CCATCAGATGTAAATGTGTGCGCATCTACAGCACAGGTATCTGTGACACCGCAACGAATGA	4082
Db	9223	CGTATGATGTCAATGTGGGCGCATCTACAAAGACCGGAACCGTTTGCACCGCAATATCA	9282
Oy	4083	ACGGCTGCACTAAAGCAAAACATGTGCTTGTGCTTCAAAGCGACCAACGAATGGGTCT	4142
Db	9283	CTGTGTTCACCGCGCATACAGCAAGTGTGCTTGTGCTTAAAGAGGACCAACGAATGGGTCT	9342
Oy	4143	TCAACTGCGCGGATCTTATTAAGGACACACACCTCAGTGCAGATTAATYAGCAATTC	4202
Db	9343	TCAACTCAGCGGATCTTATTAAGATGACACGACCAAGCGCCCAAGGGAAATYAGCAATTC	9402
Oy	4203	CATTCGCTTGAACACGACAGTGTGCGCGGTTCCGTTAGCTCACAGCGCTTCACTGACGA	4262
Db	9403	CTTTCAAGTTATCTCCAGTACTCTGCAATGTGTCTTGTGCGCACGCGCGCAATGTAAATAC	9462
Oy	4263	AGTGTTCAAAGGATCAACCTTCCACTGACTGTCAATGCGAACAACTTGTGCAACGA	4322
Db	9463	ATGTCTTTAAACATCATGAGCTTCCAATTAAGATACAGAACCTTGAACATTTGCTCACACA	9522
Oy	4323	GAAAAATTTGGGCTGTGAGACGACACCAACAGCAATGTGATTAACAGGCTTATCATCCAGGA	4382
Db	9523	GGAACTTAGGGGCAAAACCGGAAACCAACATCTGATGATGTGTGGAAAGACGCTTCAGAA	9582

QY 4383 ATTTTCTGTGGGGGAGAGAGGCTGAGTACGTATGGGGTTAACCATGAACCAAGTCAAG 4442
Db 9583 ACTTACCGCTGACCGAGATGGCCTGGAAATCATATGGGAAATCATAGCCAGTGAAG 9642
QY 4443 TCTGGGCCAGAGAGTGGGACCAAGGCAACCATGATATGGCCGATAGATCATATCC 4502
Db 9643 TCTATGCCCAAGAGTCAACACAGGAGACCTCAAGATGGGCCACAGAAATAGTACAGC 9702
QY 4503 ACTATATCATCGGATCAGTCTACATGTCAATTGTGTGTGTGTGTGTGTGTGTGTGTGT 4562
Db 9703 ATTAATACATATGCTATCTGTATACATCTTAAGCCGTGATCACTACCTGCGGGA 9762
QY 4563 TCTGTGAGGACATGATCATATGACAGCTTGATGCGCAAAAGCAAGAGAGTCTGTA 4622
Db 9763 TGAATATGGGGTAAACGTTGCAAGTATATGTGCTGTAAGCCGCGCTGATGCTGTA 9822
QY 4623 CGCCATAGCGCTTGGACCGCAACGCAAGTACCAAGCATTTAGCGGTTTGTGCTGA 4682
Db 9823 CGCCATAGCGCTTGGACCGCAACGCGTAATCCCAATCTTGTGCTGCACTTGTGTGCG 9882
QY 4683 TTGGCCCAACCAAGCGTGAACATTTGAGAACTTTAACCATCTGTGTGTGTGTGTGTGTGT 4742
Db 9883 TTAGTTCGGCCAAATGCTGAAGTTCACCGAGACCATGATTAATTTGTGTGCAAGTC 9942
QY 4743 AACCGTTTCTGGGACAGTGTGATCTCTGAGAGCGCTTGTATTTCTGTCCGCT 4802
Db 9943 AGCGGTTTCTGGGACAGTGTGATCTCTGAGAGCGCTTGTATTTCTGTGTGTGTGTGTGT 10002
QY 4803 GCTTTTCTGTGATGCTCTTTTATTTATTTGTGTGAGCGCTGTGCTGGGAAAGTGAAGC 4862
Db 10003 GCTCTCTGTGCTGCTCTTTTATTTATTTGTGTGAGCGCTGTGCTGGGAAAGTGAAGC 10062
QY 4863 CCTTGAACATGCGACCACTGTCGCAATTTTCCGGGATCCCGTATTAAGCGCTGTGTCG 4922
Db 10063 CTATGAACATGCGACCACTGTCGCAATTTTCCGGGATCCCGTATTAAGCGCTGTGTCG 10122
QY 4923 AACCGGAGTGTGAGCGGACCTTAACCTGAGATGACAGCTGTGATTTGATTTGATTTGAT 4982
Db 10123 AAAAGGAGGAGTGTGAGCGGCTCAATTTGAGATGATCTGTATTTGCTGTGAGGTTTGTG 10182
QY 4983 CTTCATCTAACAGAGTGTGAGCTGTGCAATTTGCAACAGTCAATTTTCAACCAAG 5042
Db 10183 CTTCATCTAACAGAGTGTGAGCTGTGCAATTTGCAACAGTCAATTTTCAACCAAG 10242
QY 5043 TTAATGTGTGGGCTCTGAGTGTGAGGATCTCAAGCGGATTTACATGTCGCG 5102
Db 10243 TCAATGTGTGGGCTCTGAGTGTGAGGATCTCAAGCGGATTTACATGTCGCG 10302
QY 5103 TTTTGGGAGTGTGATCCCTTTCAATGTGGGAGGCGCAATGCTTGTGCAAGTGA 5162
Db 10303 TCTTGGGAGGAGTGTGATCCCTTTATGTGGGAGGCGCAATGCTTGTGCAAGTGA 10362
QY 5163 ACACACAATGAGTGTGAGGCGTACGTCGATTCGCTGACAGTGTGCACTATAGATCAAG 5222
Db 10363 ACACACAATGAGTGTGAGGCGTACGTCGATTCGCTGACAGTGTGCACTATAGATCAAG 10422
QY 5223 TCGCACTAAAAGTTCACACAGCTGTCTGAAAGTGGCGCTGTGATTAATAGCGCAACA 5282
Db 10423 AGGGATTTAAGGTTCACACAGCTGTCTGAAAGTGGCGCTGTGATTAATAGCGCAACA 10482
QY 5283 CCACCGGCGACCTGTGATGCTTTGTCAATGTGGGCGTCAAGGTTTCCACAGGAGCGTGA 5342
Db 10483 CTACCAATTTCTTGAATGTGATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 10542
QY 5343 AGGTCTATAGCAGGCGCATATCAAGCGCTTTTTCACCTTTGACCATAGGTGTGATCA 5402
Db 10543 AAGTCTATAGCAGGCGCATATCAAGCGCTTTTTCACCTTTGACCATAGGTGTGATCA 10602
QY 5403 GAAAGGGCTGTGTATCACTACATGCTTCTGAGTATGAGCTGTGAAACAGAGCGT 5462
Db 10603 ATCGCGGCTGTGTATCACTACATGCTTCTGAGTATGAGCTGTGAAACAGAGCGT 10662

QY 5463 TCGGCAATATCAAGCACTCTGCTGTGATGCTAGACATATGATCCCGCATCTGATAC 5522
Db 10663 TTGGAGCATTTAAAGTCACTCTCTTGAATACAGAGATCTATGCGCAGACAGCATTA 10722
QY 5523 GCGTGTGAAGCTTTCTGTCAAGAACATCCAGTTCCTTACACCCAGAGATATCAGGT 5582
Db 10723 GCGTGTGAAGCTTTCTGTCAAGAACATCCAGTTCCTTACACCCAGAGATATCAGGT 10782
QY 5583 ATGAATGTGAAGAACAACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5642
Db 10783 TTGAATGTGAAGAACAACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10842
QY 5643 AAATGAATGTGAAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5702
Db 10843 AGATTTGAGTGAATTCGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10902
QY 5703 TTGAATCTCTGATGAGAGCTTTTGTGAGATCATGAAATCAACCAATTTTGAAGTTA 5762
Db 10903 TTGAATCTCTGATGAGAGCTTTTGTGAGATCATGAAATCAACCAATTTTGAAGTTA 10962
QY 5763 GCTGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5822
Db 10963 AATGTGAATGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11022
QY 5823 ACAAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5882
Db 11023 ATGTATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11082
QY 5883 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5942
Db 11083 AAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11142
QY 5943 GCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6002
Db 11143 GTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11202
QY 6003 GTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6062
Db 11203 GTAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11262
QY 6063 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6122
Db 11263 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11322
QY 6123 TCAATTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6182
Db 11323 TATTAATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11382
QY 6183 GACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6241
Db 11383 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11442
QY 6242 GCATTAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6261
Db 11443 GCATTAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11462

RESULT 14
AR096558
LOCUS AR096558 11703 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 8 from patent US 6008035.
ACCESSION AR096558
VERSION AR096558.1 GI:10025468
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11703)
AUTHORS Johnston,R.E., Davis,N.L. and Simpson,D.A.
TITLE System for the in vivo delivery and expression of heterologous
genes in the bone marrow
JOURNAL Patent: US 6008035-A 8 28-DEC-1999;
FEATURES Location/Qualifiers

source 1. .11703
/organism="unknown"
BASE COUNT 3306 a 3051 c 2913 g 2433 t
ORIGIN
Query Match 36.8%; Score 2388.4; DB 6; Length 11703;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3698; Conservative 0; Mismatches 1981; Indels 41; Gaps 8;
QY 580 ACTGACGGTATGAAGCGGAGCGGATATTTTCTCATGGAAGCGGCGCAAGTACCTTC 639
DB 5746 ACTGACCTAACCGGGGTGGTGGTACATATTTTTCAGCGACAGCGGCGGCGGACCTTC 5805
QY 640 AACGAAATGATGATGCTCAATGTAACTACAGAAACCTATATTTGGATCGGCGCTTCATG 699
DB 5806 AAAAGAAAGTCCGTTCTGACAGAACCAAGCTTACAGAACCGACCTTGGAGCGCAATCTCC 5855
QY 700 AGAAGTATACGCGCGCGCGCTGATCTGAAAGAGAAATGTACAGAAAGAACTGC 759
DB 5866 AAAGAAATTCATGCCCGGTGCTCGACACGTCGAAAGAGAAACAACCTCAAGTAC 5925
QY 760 AATTATCGGCTCTGAAGGAAATAGAACGATCATACGAAAGTAAATATGA 819
DB 5926 AGATGATGCCCAAGGAGCCAAAGAGTACAGTCTCGTAAAGTAAATACGA 5985
QY 820 AAGCAATTAACGCGGAGCACTATTTCTGGATTGGGCACTATCTATCATCAGAGTGA 879
DB 5986 AAGCATTAACCACTGAGCGCACTACTGACGATCACTGATATTAATCTTCCACAG--- 6042
QY 880 ATCTGTGAGGTGTACAGATCAATTAATCTGTACCAATCTACTCGTCAACGGTAAATTA 939
DB 6043 ATACGACGAAATGCTATTAAGATCACTACGAAACCAATTTACTCCAGTACGCTACCGG 6102
QY 940 ACAGGTTACATCTGACAGAGTCCGCGTTAAACGTGCACTTATGTTATCCAAAGAAATT 999
DB 6103 CGAATCTCTCGATCCACAGTTCGCTGTAGCTGTCTGTAAACATATCTGATAGAAACT 6162
QY 1000 ACCCTACAGTAGCAGTATATGTATTAACAGATGAATACGATCGATCTTACATGGTGG 1059
DB 6163 ATCCGACAGTACGATCTTATAGATTAAGTACGATGATGATGATCTTATGATATGGTAG 6222
QY 1060 ACGGCGCATCGTGTCTGTACAGTACAGCACTTTTTCGCGGCTAACTGAGAGCTAAC 1119
DB 6223 ACGGGAAGTCCGCTGCTGCTGATCTGCACTTCTGCGCGCTAGGTTGAAGTTAAC 6282
QY 1120 CAAGAAGCATAGCTATTTGACAGCCAGATTAAGATCAGCCGTCCCATCGCTTATACGA 1179
DB 6283 CGAAAAAACAATGAGTATAGACCCCGAATATCCGACGTGCGTTCCATCAGCGATGACA 6342
QY 1180 ATACATTAACAAATGTATGTGCTGACGATCTTAAAGGAATTTGCAACGTTCCCAATGC 1239
DB 6343 ACGAGCTACAAATGTGCTCTATTCGCGCACTAAAGAAATTTGCAAGCTCAGAGATGC 6402
QY 1240 GAGAATTAACCTGCTTGAATTCGCGGATTTAAATGTTGATTTTCAAGAAATACGAT 1299
DB 6403 GTGAATCGCCAACTGACTCAGCACTTCAATGTGATGCTTTCGAAATATATGAT 6462
QY 1300 GCATGATGATCTAGGATACCTTTGGCGATTAACCTTATTCGGCTAATCAAGAAAG 1359
DB 6463 GTAATGACGATATTGGAGAGATTGCTCGAAGCCCAATTAAGATTAACCTAAGTTTG 6522
QY 1360 TTACGCAATATGTACAAAGCTGAAAGGCGGAAAGCAGACGATTTGTTTCGAATCTC 1419
DB 6523 TCACCGCATATGTAGCTAGCTGAAAGGCGCTTAAGGCGCGCGCACTATTTTCAAAAGAG 6582
QY 1420 ATAAATCTAAACCGCTTGACAGAGATACCAATGATCAATTCGTCAATGATTAAGAG 1479
DB 6583 ATAAATTTGTGCTCAATTTGCAAGAGTGCCTATGATAGATTGCTATGAGCAATGAAGAG 6642
QY 1480 ATGTCAAGTATCTCCGCGCAGAACTATACAGAGAGCGGCTTAAGGTGAGGTTATTC 1539
DB 6643 ACGTGAAGATTACACAGGCAAGAAACACAGAGAAAGAACCGAAAGTACAAATGATGATAC 6702

QY 1540 AGGCTGAGATCCCTTGTCTACCGCTTACCTTTGCGGATCAATCGGAAATTAATCCGTA 1599
DB 6703 AAGCGGAGAACCCCTGCGACACTGCTTACTATGCGGGAATTCACCGGGAATTAATGCTGA 6762
QY 1600 GACTGAATGCGGTCTTCTGCGCAATATCAATCTCTTTCGATGATCTCAGCGGAATTT 1659
DB 6763 GGGTTACGCGGCTTCTGCTTCCAAACATTCACACGCTTTTGTGACATCTCGCGGAGATT 6822
QY 1660 TTGATGCGATTAATTTGCGAACTTTCCACAGCGGCGACAGATTTGGAAGCGACATCG 1719
DB 6823 TTGATGCGATTAATTTGCGAACTTTCCAGCAAGCGACCGGATCTGGAAGCGATATCG 6882
QY 1720 CGTCGTTTGAATAAGCGAAGACGACTATTCGCAATTTTCGGCGTGTATGATCTCTGAGG 1779
DB 6883 CATCATTTGCAAAAGCGAAGACGACTATTTGCGTTAACCGGCTGATCATCTTGTGAGG 6942
QY 1780 ACTTAGGTGACCAACCGCTCTTATGATTTGATAGGCGGCGGTTTCGCAATATCAT 1839
DB 6943 ACCGTGGTGTGATCAACCACTACTGATGATGATGATGATGATGATGATGATGAT 7002
QY 1840 CTGTGACCTTACCTACAGAAACGAGTTTAAATTTGGTGCATGATGAATTCGGTATGT 1899
DB 7003 CCACCAATCTACCTACGCGTATCTGTTTAAATTCGCGGCGATGATGAATTCGGAATGT 7062
QY 1900 TCTTAACGCTGTTTGTCAACACACTAATATATGATGATGATGATGATGATGATGATGAT 1959
DB 7063 TCTCAACATTTTGTCAACACAGTTTGAATGATGATGATGATGATGATGATGATGATGAT 7122
QY 1960 AACGTTAACAGCTCAGCGCTGCGGCTCTATTCGCGCAGATTAATATGATGATGATG 2019
DB 7123 AGCGGCTTAAACGTCAGATGTCAGCGCTTATTCGCGCAGATTAATATGATGATGATG 7182
QY 2020 TCGTCTCGACACCTTGAATGCGGAGAGATGCGGCACTTGGCTGAACATGGAATGAATA 2079
DB 7183 TAGATCTGAAGAAATATGCTGAGAGTGCACCTGCTCAATGAGATGAATGAATA 7242
QY 2080 TTATGATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2139
DB 7243 TCATGACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7302
QY 2140 ACCAGATTAACAGGACAGCTGCAAGATGCGACACCTCTTAAAGGCTTTTAACTTG 2199
DB 7303 ATTCGTTTACTTCCACAGCGCTGCGCTGCGGACCCCTGAAAGGCTTTTAAAGTTGG 7362
QY 2200 GAAACCATTTGCAAGTATGATTAACCAAGACTGCGACCGCGCGGCGCATGATGATG 2259
DB 7363 GTAAACCGCTCCAGCGGACGAGCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 7422
QY 2260 AAGCAATGCAATGAGAACAGATTTGAATTAACGACGATTAAGTAAAGCGCGTAAATCCA 2319
DB 7423 AAACAAAGGCGGTGTTAAGTATGATTAACAGGCACTTAAAGCGGTGACGACCC 7482
QY 2320 GATTCGATCATATCTGCGAGGCGCTGATCATCACTCTCTGTCAGCTTACGCTTAAAGCG 2379
DB 7483 GGTATGAGTATGACAAATATTAACCTGCTTACTGAGTATGAGAACTTTTGCACAGCA 7542
QY 2380 TTAAAGAACTTCAAGAGCATTAAGAGGAGCCCAATCACTCTTACGCTGACCTTAAATAG 2439
DB 7543 AAAGAGATTTCCAGGCAATGAGGAGGAAATTAAGCATCTCTAGGCTGCTTAAATAGT 7602
QY 2440 TGAAGTATGAGACAGCACTAC-----CCACCGGAGAAATTTTCCATACC 2486
DB 7603 CAGCATATGATCATTTTCACTGATTAATATTAACCAACACACACATGATTAAGAGATTCT 7662
QY 2487 CTCAGCTGAACCTTTCACAGCTTATACCTTACCAATCCGATGCTTACCGGATCAAAAC 2546
DB 7663 TTAAATATGCTGCGCGCGCGCTTCCGCGCGCCCACTGCAATGATGAGGCGCGGAGAA 7722
QY 2547 CTCTAGGCGCGCTGAGGCGGCTTTCGCGCGCGCGCGCGCTGCAATGGAATCTTA 2606
DB 7723 GAGGCAAGGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7782
QY 2607 GAGGCTGATATGATCACTTGAATTTCAACCAAGATCACTTAATCCGCGCGAGTCCA- 2665

Db 7783 CCACAGCCGTAGTCCCTAGTCATTGGACAGGCAACTAGACCTCAACCCCACTGCAC 7842
Qy 2666 -----CCGCCAAGAGAGAGAGAGTGTCTTAACCCAAAACCTACTGACCTTAAAGA 2720
Db 7843 GCCCGCCACCGCGCCAGAGAGAGAGGCGCCCAACCAACACACGAGCCGAGAAACCA 7902
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGAGAAAC---GCAAGCTTAAACAGAGGAAACGACACGTA 2777
Db 7903 AAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7962
Qy 2778 TGTGTATGAGTGTGAGTGTGAGCAAGACATTTCCGATCATG---CTGAACGCGCCAGTGA 2834
Db 7963 TGGCATTAATGTTGAGAGCCGACAGATTTGTCAGTCAGAGAAAGAGAGAGAGATGCA 8022
Qy 2835 ATGATATGCTGCGCTTTCGAGAGAGGCTGATGAACCACTCCAGCTTGAAGAAAA 2894
Db 8023 TCGGCGACGACTGCGCATGAGAGAGAGATTAATAACCTTCGACGTTGAAGAGACCA 8082
Qy 2895 TTGATATGAGCAATTAGCGCGCTGAATTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2954
Db 8083 TCGACACCCCTGTGTATCAAAAGCTCAATTTACCAAGTCGTGACATACAGATGAGT 8142
Qy 2955 ACGGCGAGCTTCCCGAAGACATGAATCAGACACGCTGCAAGTCAACAGCAAAACAC 3014
Db 8143 TCGCAGATGCGACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8202
Qy 3015 CGGCGTTTACACATGCGACCAACGCGCGAGTCAGTATGAGAGATGAGAGATTTACCGTAC 3074
Db 8203 AAGGATTTATTAATCGGACCAACGAGGCGGTGACATATGAGAGATTTACATCC 8262
Qy 3075 CGAGAGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3134
Db 8263 CTCGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8322
Qy 3135 TTGTGCTATTTGTTCTAGAGAGTGCATTAAGGAGAGAGAGAGAGAGAGAGAGAGAG 3194
Db 8323 TTGTGCGATATGTCCTCGGTGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 8382
Qy 3195 CTTGGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3251
Db 8383 CTTGGAATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGCTGTTCGATGTCAGTTCCTCATGCGCAAC 3305
Db 8443 CAGACACACTGCTGACGCGCATATGTTTGTCTCGAGAGAGAGAGAGAGAGAGAGAG 8502
Qy 3306 CACCGTGTGCTATTTCACTGACGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
Db 8503 CGCCACATGCTATACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8562
Qy 3366 ACAATCCAAATTAAGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
Db 8563 ACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8622
Qy 3423 CCAAGCAGAGATTAACGATGATCACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3482
Db 8623 GCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8682
Qy 3483 ATTGAGACACTCAACGCGTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3542
Db 8683 ACTGACACATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8742
Qy 3543 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3602
Db 8743 CGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8802
Qy 3603 CTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3662
Db 8803 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8862
Qy 3663 ACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3722

Db 8863 GCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8922
Qy 3723 GGTACTTCTGTTAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3782
Db 8923 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8982
Qy 3783 GAGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3842
Db 8983 GCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9042
Qy 3843 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3902
Db 9043 AAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9102
Qy 3903 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3962
Db 9103 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9162
Qy 3963 CCTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4022
Db 9163 CCTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9222
Qy 4023 CCTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4082
Db 9223 CGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9282
Qy 4083 ACGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4142
Db 9283 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9342
Qy 4143 TCAATCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4202
Db 9343 TCAATCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9402
Qy 4203 CATTCGCTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4262
Db 9403 CTTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9462
Qy 4263 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4322
Db 9463 ATGCTTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9522
Qy 4323 GAAATTTGGGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4382
Db 9523 GGAATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9582
Qy 4383 ATTTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4442
Db 9583 ACTTACCGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9642
Qy 4443 TGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4502
Db 9643 TCTATGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9702
Qy 4503 ACTATTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4562
Db 9703 ATTTACTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9762
Qy 4563 TCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4622
Db 9763 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9822
Qy 4623 CGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4682
Db 9823 CGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9882
Qy 4683 TTGCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4742
Db 9883 TTAAGTGGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9942
Qy 4743 AACCGTTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4802
Db 9943 AGCGGTTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10002

Qy	4803	GCCTTTACATGCTGATGCGCTTTTATTTATGCTGTCCAGGCGCTGCGGGAAAGGTAGAG	4862
Db	10003	GCTGCTCTCTGCTGCTGCTCTTTTATGATGCTTCCGGCGCTACCTGGGAAAGGTAGAG	10062
Qy	4863	CCTTCGAAACATGCGACCACTGTGCGCAAAATGTTCCGGGAAATCCCGTATTAAGCGTTGTCG	4922
Db	10063	CCTACGAAACATGCGACCACTGTTTCCAAATGTGCGACGAATACCGTATTAAGCACTTGTTG	10122
Qy	4923	AAGCGGAGGTTACGGCGCACTTAACCTGGAATTCACGGTGGTCTATGCGAAATTAACAC	4982
Db	10123	AAAGGGGAGGGTATGCCCCGCTCAATTTGGAAGATCACTGTCAATGTCCTGGAGGTTTTCG	10182
Qy	4983	CTTCAACTAACGAAGATAGTGACCTGACCTGCAAAATTCACACAGTCATTTCTTACCAACAAG	5042
Db	10183	CTTCCACCAACCAAGATATTAATCTGCAAAATTCACACTGTGGTCCCTCCCAAAA	10242
Qy	5043	TTAAATGCTCGGGTCCCTCGAGTCGAAGGATCTCAAAAGCGGATTAACATGCGCGG	5102
Db	10243	TCAATATCTCGGCTCCTTGGAAATGTACGCGCGCCCTCATGACAGCTATACCTGCAAGG	10302
Qy	5103	TTTTTGGGAGTGTATCCCTTTATATGTGGGAGGCGCAAAATGCTTCTGTGACAGTGA	5162
Db	10303	TCTTCGAGAGGCTTACCCCTTTATGTGGGAGGAGGCAATGTTTTCGACAGTGA	10362
Qy	5163	ACACACAACTAGTGAAGCGGTACGTGAGTTCGCTCGACTGCACTATAGATTCACGACG	5222
Db	10363	ACACACAGATGAGTGAAGCGGTACGTGAACTGTCAAGATTTGCGGCTGTGACACGCGCG	10422
Qy	5223	TCCGACTTAAAGTTTCAACACAGCTGTCTGAAATTCGGCCCTGCGTATAGTATACGCACA	5282
Db	10423	AGGCGATTAAGTGTCACTACGCGCGCATGAAATGAGATCTGCTATAGTATCGGGAACA	10482
Qy	5283	CCAACGCGCACTGATACGTTTGTCAATGCGCTCAAGCGAGTTTCTCACGGGACTGA	5342
Db	10483	CTACCACTTTCTTAGATGTGTAGTGAAGGATTCACACAGGAAGCTTAAAGACTTGA	10542
Qy	5343	AGGTCAATGACAGGCGCGATATCACGCGCTTTTCAACCCTTTGACCAATTAAGTGTGATCA	5402
Db	10543	AAGTCATAGCTGACCACTATTTCAAGCATCTGTTACGCCATTCGATCAATTAAGTGTGATTC	10602
Qy	5403	GAAAGGGGCTTGTTTTCAACTAGACACTCCCTAGATATGAGTATGAATACCAAGGAGCT	5462
Db	10603	ATCGCGGCTGTGTGTACACTATGACTTCCCGAATATGAGGAGTGAATCAAGAGGCT	10662
Qy	5463	TCGGCGATTTTCAAGCATCTCGCTTGAATGCTACAGACATAGTAGCCCGCACTGACATAC	5522
Db	10663	TTGGAGACATTTCAAGCATCTCTTGTGACTAGCAAGATCTCATCGCCAGCACAGACATTA	10722
Qy	5523	GGCTGCTGAAGCCTTCTGTCAAGAACATCCAGTCCCCCTTACACCAAGCATATCAGGCT	5582
Db	10723	GGCTACTCAAGCCTTTCGCGAAGAACGTGTATCCCGTACAGCAGGCGGATCAAGAT	10782
Qy	5583	ATGAAATGTGGAAGAAACAATCACTAGAGACGCCCTGCAAGAAACAGACCAATTTGATGTA	5642
Db	10783	TTGAGATGTGGAATAAACAATCAAGCGCGCCCACTGACAGAAACCGCACTTTTCGGGTGTA	10842
Qy	5643	AAATTTGAATGAGACCTCTGCGAGCGCTTACTAGTGTGTTACGSGCACATCCCTATCTGA	5702
Db	10843	AGATTTGACATTAATCCCTCTCGAGCGGTGACGTGTTCAATCGGGAACATTTCCCATTTCTA	10902
Qy	5703	TTGACATCCCTGATATGACGTTTTGTGATCATCAGATCACCACAAATTTTATGAAGTTA	5762
Db	10903	TTGACATCCCGAAGCGTGTCTTTATCAGACATCAAGATGACCACTGTGTCTCAACAGTGA	10962
Qy	5763	GCTGCACAGTAGCAGACTGATTTATTTCTGACAGACTTTGGTGGTTCTTAACATTAACGT	5822
Db	10963	AAATGTGAAGTACAGATGTCACTTATTCAGCAACATTCGCGGGAGATGCGCACCTGTGAGT	11022
Qy	5823	ACAAAGCTGACAGGAGGACATTTGTCCAGTTATCTCCACTCCACAGACAGCTGTTTGA	5882
Db	11023	ATGTATCCGACCGGAGAGTCAATGCCCCGTATATTTGCACTTGACACACAGCAACTCTTC	11082

QY	5883	GGNAGGAGCACCACGATGTAAGTCGCGTAGGACAGATPAACATPAKATTTTATAGACATATCGA	5942
Db	11083	AAGAGTGCACGATACATATCTCTGGAGAAAGACGGGTGACATGACATCTTTATAGACACGGCGA	11144
QY	5943	GCCCAACAGCAATATTTATAGATTTTCGCTATGCGCGCAAGAAAGTCCACCTGCATGCTGAAT	6002
Db	11143	GTCCACAGCGCACTTTATATGTAATCGCTGTGTGGAAAGAAAGAAATGCAATGCAAGAAAT	11200
QY	6003	GTAACACACCGGCGGACCAACATPAATTTGGAGAACCACTAAAGTCGACCAAGAAATTCAGG	6062
Db	11203	GTAACACACACGCTGACCAATATCGTGGACACCCGCGCAAAAATATGACCAAGAAATTTCAAG	11267
QY	6063	CGGCAATTTCCAAACATCTTTGGAACGTGGTCTGTGACATGTTTGGGGGAGATCAATCCC	6122
Db	11263	CCGCAATCTCAAAAACATCAATGAGTTGGTCTGTTTGGCTTTTTCGGGGGGCGCTTCGTGCG	11322
QY	6123	TCATTTGTTGTAGACTATATAGTGTGTGTCAGCTTATGCTTATPAACACACAGTAGAT	6182
Db	11323	TATTTATTTATATAGACTTATATGATTTTTCCTTTCGACAGCATGATGCTGACATGACACAGAAAT	11382
QY	6183	GACTGAGCGCGGACACTGA-CATPAGCGTAAAACTCGATGTACTTCGAGGAAGCGTGT	6241
Db	11383	GACCGTACGCGCCCAATGATCCGACACAGCAAAATCTGATGTACTTCGAGGAAGCATGTGT	11444
QY	6242	GCATPAATGCCACGGCGCGCT	6261
Db	11443	GCATPAATGCATCAGGCTGT	11462

LOCUS	BD071123	11703 bp	DNA	linear	PAT 27-AUG-2001
DEFINITION	System for the in vivo delivery and expression of heterologous genes in the bone marrow.				
ACCESSION	BD071123				
VERSION	BD071123.1	GI:22616726			
KEYWORDS	JP 2001515348-A/4.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 11703)				
AUTHORS	Johnston,R.E., Davis,N.L. and Simpson,D.A.				
TITLE	System for the in vivo delivery and expression of heterologous genes in the bone marrow				
JOURNAL	Patent: JP 2001515348-A 4 18-SEP-2001; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL				
COMMENT	OS Girdwood S.A. PN JP 2001515348-A/4 PD 18-SEP-2001 PF 18-FEB-1998 JP 1998536760 PR 19-FEB-1997 US 08/801263 PI ROBERT E JOHNSTON, NANCY L DAVIS, DENNIS A SIMPSON PC CI:NI5/86, CI:NI5/33, CI:NI7/01, CI:NI5/10, A61K39/12, A61K48/00 CC CC the for the in vivo delivery and expression of heterologous genes in CC the bone marrow FH Key FT source FT 1. 11703 /organism='Girdwood S.A.'				
FEATURES	Location/Qualifiers				
source	1. 11703 /organism='unidentified' /mol_type='genomic DNA' /db_xref='taxon:32644'				
BASE COUNT	3306 a	3051 c	2913 g	2433 t	
ORIGIN					
Query Match	36.8%; Score 2388.4; DB 6; Length 11703;				
Best Local Similarity	64.7%; Pred. NO. 0;				
Matches 3398;	0; Mismatches 1981; Indels 41; Gaps 8				
580	ACTGACGGTATGAAAGCGGACGGTATTTTCTCATCGGAACAGGCCAACGTCACCTTC 639				

Dh 5746 ACTGACTAACCGGGGTAGTGGGTACATATTTTTCGACGACACAGGCCCTGGGCACTTGC 5805
Qy 640 AACAGAAATCAGTACGTCAATGTAACTAACAGAACTATTTGGATCGGGCGTCCATG 699
Db 5806 AAAAGAAAGTCCGTTTTCGAGAACCAAGCTTACGAAACCACTTGGAGGCGCAATGTCCTGG 5865
Qy 700 AGAATATTACGCCCCCGCGCTCGATCTCGAAAGAGAAAATGTTACAGAAATACTGC 759
Db 5866 AAAAGAAATCATGCCCCGCGTCTCGACAGCTCGAAAGAGAAACACTCAACTCAGATACC 5925
Qy 760 AATTATGGCCCTCTTAAGAAATAGAGCAGGTATCAATCAGAAAAGTAAATATAGA 819
Db 5926 AGATGATCCCAACCGAAGCAAAAGTAGGTACAGTCTGTAAAGTAGAATAATCTGA 5985
Qy 820 AAGCAATTACAGCGGAGGCACTATTTCTGATTTGGGCAATATCATCATCAGAAATGA 879
Db 5986 AAGCCTAACCACTAGAGCACTACTGTACAGGACTACGACTGTATTAACCTTGCACAG--- 6042
Qy 880 ATCTGTGAGTGTACAGAGTCAATTTCTGTACCAATCTACTCTGTCACAGTAAATTA 939
Db 6043 ATCAGCCAGAAATGCTATTAAGATCACCTATCCGAAACANTGTACTCCAGTACGATCCGG 6102
Qy 940 ACAGGTTTACATCTGCAAGAGTTCGGCGTTAAAACTGCACTTAGTTATCCAGAAAT 999
Db 6103 CGAATCTCCGATCCACAGTTCGCTGTAGCTGTCTGTAACTAATCTGCATGGAATCT 6162
Qy 1000 ACCCTACAGTACCGCAGTTATTTGTATACAGATGAATAGATCGTATCTGACATGTGG 1059
Db 6163 ATCCGACAGTACGATCTTATTCAGATTTACTGACGAGATGAGATCTTACTTGGATATGGAG 6222
Qy 1060 ACGGGCATCTGCTGTCTAGATACAGCCACTTTTGTCCGGCTAACTAGAGAACTACC 1119
Db 6223 ACGGAGAGTGGCGCTCGCTGATACCTGCAACTTCTGCGCCGCTAAGTTAGAAATTTACC 6282
Qy 1120 CAAGAAGCATAGCTATTTTTCAGCCAGAGATTAAGTACAGCGTCCATCGCTTATACGA 1179
Db 6283 CGAAAAAATGATGATAGAGCCCGCAATATCCGAGTGGCTGTCAATCAGAGATGACGA 6342
Qy 1180 ATACATTAACAATAATGTATGGCTGACAGTACTAAAGAAATTTGCAACGTTCCCAAAATGC 1239
Db 6343 ACACGCTACAAATATGCTCATTTGCGGCACTTAAAGAAATTTGCAACGTCACGACGATGC 6402
Qy 1240 GAGAAATTAATCTGTCTAGATTCGGCGGATTTAATGTGATTTGTTCAAGAAATACGAT 1299
Db 6403 GTGAATCTCCAAACATGAGCTCAGGACATTTCAATGTGAAATGCTTTGGAAATATGAT 6462
Qy 1300 GCAATGATGATCTGAGATACCTTTCCGATTAACCTTATTCGCTTAACTACAGAAACG 1359
Db 6463 GTAATGACGAGTATTTGGAGAGAGTTCCGCTCGGAGCCCAATTAGATTTCACTGAGTTGG 6522
Qy 1360 TTACGCAATATGTGACAAAGCTGAAGGGCCGAAGGCGACGACATTTTGGGATATCTC 1419
Db 6523 TCACCGCATATGTAGCTAGACTGAAGGCGCTTAAAGCGCGCACTATTTTGGAAAGAGT 6582
Qy 1420 ATAAATCTAAAAACCGTTGAGAGAGATACCAATGATCAATTCGTGACATGCTTAAAGAG 1479
Db 6583 ATTAATTTGTCCCATTTGCAAGAAAGTCCATGATGATGATTCGTCAATGACATGAAAGAG 6642
Qy 1480 ATGTCAAGTTATCTCCCGGACGAAACATACAGAGAGCGGCTTAAAGGTGCAAGTTATTC 1539
Db 6643 ACGTCAAAAGTTATCAACAGGACGAAACACAGAAAGAAAGCAAGTACAGATAC 6702
Qy 1540 AGGCGCAGATCCCTTGTCTACCGGTTACTCTTTGGGGATCCATGGGAATTAATGTCGTA 1599
Db 6703 AAGCGCAGAACCCCTGGCGATGCTTACTTATGCGGATTTACCGGGAATTAATGTCGTA 6762
Qy 1600 GACTGAATGCGGTGCTTTCGCAAAATATCAATCTCTTGACATGTCAGCGGAGAAAT 1659
Db 6763 GGCCTTAAGCGCGTCTTGTCTTCAAAATTTCAACACCTTTTGAACATGTCGCGGAGAAAT 6822
Qy 1660 TTGATGCAATTAATTTGTAACATTTTCAACAGCGGACCCAGTATTTGAAACGACATCG 1719
Db 6823 TTGATGCAATCAATGACAAACCTTCAAGCAAGCGGACCCGCTACTGAGAGAGGATATCG 6882

Qy 1720 CGTGGTTGATAAAAAGCAAGACGACGCTATCGCAATTTGCGGCTTGTGATGATCCTTGAG 1779
Db 6883 CATCATTTGACAAAAAGCCAAAGACGACGCTATGCTTAAACGGTGTGATGATCTTGGAG 6942
Qy 1780 ACTTAGTGTGACCAACCGCTCTTGAATTTGATAGAGCGCGCTTCCGCAATATCAAT 1839
Db 6943 ACTTGGGTGTGATCAACCACTACTGACTGATGAGTGGCGCTTTGGAGAAATATCAT 7002
Qy 1840 CTGTGACCTTACCTTACAGAAACGAGTTTAAATTTGGTCCATGATGAATCCGTAAT 1899
Db 7003 CCACCATCTACCTTACGAGGTACTGTTTAAATTCGGGCGATGATGAATCCGAAATGT 7062
Qy 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTGACAGATACGTT 1959
Db 7063 TCTTCACTTTTGTCAACACAGTTTGAATGTCATGATGCTGACGACGAGTACTAGAG 7122
Qy 1960 AACGTTAACACGTCAGCGGTGCGCGCTCTATTCGCGAGATTAACATATGATGATGTT 2019
Db 7123 AGCGGCTTAAACGTCAGATGTGACGCTTATGCGAGCAACATCATATGAG 7182
Qy 2020 TCGTCTCGACACCTTGAATGCGGAGATGCGCACTTGGCTGAACATGGAATTAATA 2079
Db 7183 TAGTATCTGACAAAGAAATGCTGAGAGTGCAGCCTGCTCAACATGAGGTTAAGA 7242
Qy 2080 TTATGATGACGATTAATGTATCAAGACCCCTTCTGTGGGGGATTTATCCGTGTG 2139
Db 7243 TCATGACGACGATCATCGGTAGAGACCACTTACTTTCGGCGGATTTATCTTGACAG 7302
Qy 2140 ACAGATTAACAGGACGACCTTCAGAGTTCGAGACCTCTTAAAGAGCTTTTAAAGCTTG 2199
Db 7303 ATTGCGTACTTCCACAGCGGTGCGCGGTGGAGACCCCTGAAAGAGCTGTTAAAGTTGG 7362
Qy 2200 GAAACACTTGGCAGTGCATGATATACCAAGCTGCGACCGCGGCGACTGCTGATG 2259
Db 7363 GTAAACCCCTCCAGCCGACGACGAGCAAGACGAAAGACGCGCTCTGCTAGATG 7422
Qy 2260 AAGCAATGCAATGGAACGAATTTGGAATTAGGAGAGATTTGTAAGGCGCGGATTAACA 2319
Db 7423 AAACAAAGCGGTGTTTGAATGATTAACAGGCACTTTAGCACTGACGCGGTGACGACCC 7482
Qy 2320 GATACGATCATACTGACGACCGCTGATCATCAGCTCTCTGTCCAAGTTAGCCGAAACG 2379
Db 7483 GGTATGAGTATGACAAATTTATACCTGTCTTACAGCTTATGAGAACTTTGCCAGACGA 7542
Qy 2380 TTAAAGACTTAAGACATTAAGAGGAGCCCAATCACTCTTACGCGTGAACCTAAATAG 2439
Db 7543 AAAGGACATTCAGAGCCATCAGAGGGGAAATTAAGCATCTCTACGCTGCTCTAAATAGT 7602
Qy 2440 TGACGTGTAGACACGACCTAC-----CCACCGGACGAAATGTTTCCATACC 2486
Db 7603 CAGCATATGATATTTTATCTGATCTAATATCTAACACACACCACTGAATGAGATTC 7662
Qy 2487 CTCAGCTGAACCTTTCCACAGATTTTACCTTACAAATCCGATGCTTACCGAGATCCAAAC 2546
Db 7663 TTAAATGCTGCGCGCGCGCGCTTCCCGGCGCCACATGCAATGAGAGCGCGGAGAA 7722
Qy 2547 CTCTAGCGCGCTGAGCGCTTGGCGCGCTTGGCGCTGCTGCTCAATTCGAATCTTA 2606
Db 7723 GAGGACAGCGCGCGCGCTGCTGCGCAACGAGGCTGCTTCTCAAAATCCAGCACTGA 7782
Qy 2607 GGAGGTGATATGATCACTTGAACCTTCAACAAACATCACTCAATCCGCGCGAGTCTCA- 2665
Db 7783 CCACAGCGGTATGCTTCAATGATGACAGGCACTGACCTTAACCCCAAGTCCAC 7842
Qy 2666 -----CCGCAAAAGAAAGAGAGTGTCTCTTAAAGCCAAACCTTACAGGCTTAAAGAA 2720
Db 7843 GCCCGCAACCGCGCGCAAGAAAGAGAGCGCGCCAAAGCAACACGAAAGCGAAAGAAACAA 7902
Qy 2721 AGAAGCAAGCAAGCAAGAGAGCAAC--GCAAGCTTAAACAGGAAACGACACGTA 2777
Db 7903 AAAGCAGAGAAAGAAAGAAAGCAACCTGCAAAACCCAAACCGGAAAGAGACAGCGCA 7962

Db 10123 AAGGGCAGGGTATGCCCCGCTCAATTGGAGATCATGTCTATGCTCGAGGTTTTGC 10182
Qy 4983 CTTCAACTACAGAGAGTAGCGACCTGCAATTCCACAGATCTTCCTTACACAG 5042
Db 10183 CTTCCACCAACCAAGAGTAGCATTTACTCTCAATTACACACTGTGGTCCCTCCCAAAA 10242
Qy 5043 TTAAATGCTGGGGTCCCTCGAGTGCAGAGGCACTCTCAAGCGGATTTACAATGCCG 5102
Db 10243 TCAATGCTGGGGTCCCTGGAAATGTACGCGGGCGCTCATGACAGACTATACCTGCAAG 10302
Qy 5103 TTTTGGGGGTGTGATCCCTTTCATGTGGGAGGGGCAATGCTTCTGTGACAGTGA 5162
Db 10303 TCTTGGAGGGGTATCCCTTTATGTGGGAGGCGCAATGTTTGGCCACAGTGA 10362
Qy 5163 ACACACAATGAGTAGGCGTAGCTGAGTTGCTCCAGACTGCATATAGATCAGCAG 5222
Db 10363 ACAGCCAGATAGTAGGCGTAGCTGAGTTGCTCCAGACTGCATATAGATCAGCAG 10422
Qy 5223 TCGCACTAAAGTTTACACAGCTGCTGAAAGTGGCGCTGATATAGTATACGCAACA 5282
Db 10423 AGGCGATTAAGGTGACACTGCGCGATGAAGTAGAGCTGGTATATGATACGGAAACA 10482
Qy 5283 CCACCGCGCAGCTGATACGTTTGTCAATGGCGTACGCCAGGTTCTTACGGGACTGA 5342
Db 10483 CTACCAAGTTTCTAGATGTATGACGAAAGGATCACACAGAACTTAAGACTTGA 10542
Qy 5343 AGGTATAGAGGGCGGATATCAGCGGTTTTTCAACCTTTGACCATTAAGTTCGTCATCA 5402
Db 10543 AAGTATATAGCTGAGCAATTTACAGATGTTTACGCCATGATATTAAGTTCGTTATCC 10602
Qy 5403 GAAAGGGCTTTTTCACACTACGACTTCCCTGATATGAGCTATGAACAGAGCGT 5462
Db 10603 ATCCGGCGCTGGTGTACAACTATGACTTCCCGAATATGAGCGATGAACAGAGCGT 10662
Qy 5463 TCGCGATATTTCAAGCATCTCGCTGTATGCTACAGACATAGTACCCGCACTGACATAC 5522
Db 10663 TTGAGACATTTCAAGCTCTCTTGACTACAGAGATCTCATGCGCAGACACATTA 10722
Qy 5523 GGTGCTGAAGCTTTCTGTCAAGAACATCCAGTCCCTACACCCAGACATATCAGGCT 5582
Db 10723 GGTACTCAAGCTTTCCGCCAAGACGTGATGTCCTGACACGAGCGCGCATCAGAT 10782
Qy 5583 ATGAATGTGAGAGAACACTCAGGACGACCCCTGCAAGAAACAGACATTTGGATGTA 5642
Db 10783 TTGAGATGTGAGAAACAACACTCAGGCGGCCCACTGCAAGAAACGCACTTTCCGGTGT 10842
Qy 5643 AAATTGAATGAGGCTCTGCGAGGCTTAATCTGTCTTACGGGCAATCCCTATCTGA 5702
Db 10843 AGATTGCGATTAATCCGCTCCGAGCGGTGAATGTTCAATCGGGAACATTCCTATCTTA 10902
Qy 5703 TTGACATCCCTGAGCGCTTTGTGAGATGATCAGATTCACCAAACTTTAGAGTTA 5762
Db 10903 TTGACATCCCTGAGCGCTTTGTGAGATGATCAGATTCACCAAACTTTAGAGTTA 10962
Qy 5763 GCTGCACAGTAGACACTGATTTATTTCTGAGACTTTGTGTCTCTTAACATTACAGT 5822
Db 10963 AATGTGAAGTAGTAGAGTGCATTTATGACAGACTTGGCGGATGGCCACCTGCAAGT 11022
Qy 5823 ACAAGCTGACAGGAGGAGCAATGTCTCACTTCACTCCACTCCAGACAGCTGTTTTGA 5882
Db 11023 ATGTATCCGACCGGAGAGTCAATGCCCCGTACATTCGATTCGAGCACACAACTCTCC 11082
Qy 5883 AGGAAGGACCAACATGTGACTGCCGATAGCAGATTAACACTATTTTAGCATGTA 5942
Db 11083 AAGAGTGACAGTACATGTCTTGAGAAAGAGCGGTGACAGTACATTTAGCACCGGA 11142
Qy 5943 GCCCAAGCAAAATTTTATGTTTGTATGCTATGCGCAAGAACTCCAGTCAATGCTGAAT 6002
Db 11143 GTCCACAGGCGAATTTATGCTATGCTGTGTGGAGAAAGACAACTGCAATGCAAGAT 11202
Qy 6003 GTAAACCAACCGCGGACCACTTAATTGGAACCACTAAAGTGCAGCCAAAGATTCAGG 6062
Db 11203 GTAAACCAACAGTGCATATGAGCAACCCGCAAAATGACCAAGATTTTCAAG 11262

Qy 6063 CGGCACTTCCAAAACATCTTGAAGTGGCTGTGCACTGTTTGGGGAGCATATCC 6122
Db 11263 CCGCATCTCAAAAACATCATAGAGTTGGTGTGCTTTTGGGGCGCCCTGCTGC 11322
Qy 6123 TCATTTGTAGAGCTTATAGTGTGTGACCTCTATGCTTATTAACACAGTAGAT 6182
Db 11323 TATTAATTAATGACTTATGATTTTGTCTGACAGATGATGCTAGTACACAGAGAT 11382
Qy 6183 GACTGACGGGAGACCTGA-CATAGCGTAAACTCGATGACTTCCGAGGAGCGTGT 6241
Db 11383 GACGCTACGCCCAATGATTCGACAGCAAACTCGATGATCTCCGAGGAGCTGATGT 11442
Qy 6242 GCATPATGCCACGGCGCT 6261
Db 11443 GCATPATGCATCAGCTGTGT 11462

Search completed: November 16, 2003, 05:39:13
Job time : 22826 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 12:03:02 ; Search time 8996 Seconds
(without alignments)
13508.491 Million cell updates/sec

Title: US-10-023-649-1_COPY_1_5000

Perfect score: 5000

Sequence: 1 accatcaacaatacgaac.....ataatctagaagccacc 5000

Scoring table: IDENTITY NUC
Gapox 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rtd:*
26: em_ges_phg:*
27: em_ges_vrt1:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match length	ID	Description
C 1	49.8	1.0	885 13 BX425603	BX425603 BX425603
C 2	47	0.9	1201 13 BX381961	BX381961 BX381961
C 3	43.8	0.9	1098 13 BX436460	BX436460 BX436460
C 4	42	0.8	1201 13 BX381961	BX381961 BX381961

C	5	41.2	0.8	646 14	CD393733
C	6	40.8	0.8 <td>398 12</td> <td>BG984025</td>	398 12	BG984025
C	7	40.8	0.8 <td>971 13</td> <td>BUI73419</td>	971 13	BUI73419
C	8	40.8	0.8 <td>1101 29</td> <td>CNS0175Y</td>	1101 29	CNS0175Y
C	9	40.6	0.8 <td>380 12</td> <td>BUI98034</td>	380 12	BUI98034
C	10	40.6	0.8 <td>453 12</td> <td>BUI98451</td>	453 12	BUI98451
C	11	40.6	0.8 <td>455 12</td> <td>BUI98768</td>	455 12	BUI98768
C	12	40.6	0.8 <td>456 12</td> <td>BUI98768</td>	456 12	BUI98768
C	13	40.4	0.8 <td>715 12</td> <td>BUI95766</td>	715 12	BUI95766
C	14	40.4	0.8 <td>1201 13</td> <td>BX376097</td>	1201 13	BX376097
C	15	40.2	0.8 <td>634 10</td> <td>BE017217</td>	634 10	BE017217
C	16	39.8	0.8 <td>413 14</td> <td>CB384158</td>	413 14	CB384158
C	17	39.8	0.8 <td>916 10</td> <td>BG541362</td>	916 10	BG541362
C	18	39.6	0.8 <td>343 9</td> <td>AMS01797</td>	343 9	AMS01797
C	19	39.6	0.8 <td>402 9</td> <td>AMS01918</td>	402 9	AMS01918
C	20	39.6	0.8 <td>449 9</td> <td>AM499678</td>	449 9	AM499678
C	21	39.6	0.8 <td>480 9</td> <td>AM499647</td>	480 9	AM499647
C	22	39.6	0.8 <td>491 9</td> <td>AM502451</td>	491 9	AM502451
C	23	39.6	0.8 <td>859 14</td> <td>CA971504</td>	859 14	CA971504
C	24	39.6	0.8 <td>1011 10</td> <td>BE563150</td>	1011 10	BE563150
C	25	39.4	0.8 <td>504 13</td> <td>BUI926381</td>	504 13	BUI926381
C	26	39.4	0.8 <td>597 12</td> <td>BM308456</td>	597 12	BM308456
C	27	39.4	0.8 <td>619 13</td> <td>BUI927494</td>	619 13	BUI927494
C	28	39.2	0.8 <td>999 12</td> <td>BM477724</td>	999 12	BM477724
C	29	39	0.8 <td>243 14</td> <td>CB391641</td>	243 14	CB391641
C	30	39	0.8 <td>560 28</td> <td>AZ738168</td>	560 28	AZ738168
C	31	39	0.8 <td>599 13</td> <td>BQ873224</td>	599 13	BQ873224
C	32	39	0.8 <td>626 13</td> <td>BQ849532</td>	626 13	BQ849532
C	33	39	0.8 <td>997 29</td> <td>CNS0057E</td>	997 29	CNS0057E
C	34	39	0.8 <td>1240 10</td> <td>BG739729</td>	1240 10	BG739729
C	35	38.8	0.8 <td>342 12</td> <td>B1177442</td>	342 12	B1177442
C	36	38.8	0.8 <td>650 28</td> <td>BH900647</td>	650 28	BH900647
C	37	38.8	0.8 <td>687 12</td> <td>BM406647</td>	687 12	BM406647
C	38	38.8	0.8 <td>705 12</td> <td>B1432608</td>	705 12	B1432608
C	39	38.8	0.8 <td>718 10</td> <td>BG598366</td>	718 10	BG598366
C	40	38.8	0.8 <td>1101 29</td> <td>CNS017GS</td>	1101 29	CNS017GS
C	41	38.8	0.8 <td>1182 9</td> <td>AL513775</td>	1182 9	AL513775
C	42	38.4	0.8 <td>413 13</td> <td>BQ281915</td>	413 13	BQ281915
C	43	38.4	0.8 <td>1003 29</td> <td>CNS04H1P</td>	1003 29	CNS04H1P
C	44	38.4	0.8 <td>1076 29</td> <td>CNS05HXN</td>	1076 29	CNS05HXN
C	45	38.4	0.8 <td>1163 29</td> <td>CC290181</td>	1163 29	CC290181

ALIGNMENTS

RESULT 1
LOCUS BX425603 885 bp mRNA linear EST 15-MAY-2003
DEFINITION BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens CDNA clone
CLOBB022ZA07 3-PRIME, mRNA sequence.

ACCESSION BX425603
VERSION BX425603.1 GI:30770486

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 885)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization

COMMENT Unpublished

CONTACT: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@life.techn.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradey Avenue Genoscope sequence ID : CLOBB022ZA07FP1.

Location/Qualifiers

1. .885

/organism="Homo sapiens"

BASE COUNT	206 a	134 c	134 g	172 t	host cells (Stratagene)."
ORIGIN					
Query Match		0.8%	Score 41.2;	DB 14;	Length 646;
Best Local Similarity		48.9%	Pred. No. 3.4;		
Matches 139;	Conservative	0;	Mismatches 143;	Indels 2;	Gaps 1;
QY	2331	AGTGCCTTAAATGGGGTTAGACCCCGTTAACACTCTGTACATTGATGAGCATTTGC	2390		
Db	599	AGTCTTCGCAATTCAGAGGGGTTTAAAGATCATTAGTGTGTGTTTCCAAACATGCTAA	540		
QY	2391	CTGCCATGCAAGGAGCGCTGTGCATCATGATGCCATCGTCAAACTTAAGAAAGTGTAAT	2450		
Db	539	GTTCCTTTGCTGCTATGATCTTTGGCAGACTTTTGGCATTTCCATCTTCAAAACAGTGTAAT	480		
QY	2451	GTGGGGGAGCCCAAAACAATGGCGCTCTTTAACATGATGCTGAAGATTAATTTAA	2510		
Db	479	GCGCTCTTAATGAACAGAGAGGCCACTGTATGACACAGACAAATATGAAAGTGAAG	420		
QY	2511	CCATGACATATGCACTGAAGTGAACCATTAACAATCTCTAGAGAGTGCAACAGACTGT	2570		
Db	419	TGTCAAGATTAACGTTAAATGCCCCATCACAATG--CTCAGAGGAGTGTAGCCAGACCA	362		
QY	2571	AACCGCATGTCTCCACGCTCTTTTACGACAAAGCCGATGAAG	2614		
Db	361	ATACCAACAGCCCTGCATGTTTCTTGTGCAGAAAGCTGCAGAG	318		
RESULT 6					
LOCUS	BC984025	398 bp	mRNA	linear	EST 12-JUN-2001
DEFINITION	H15-CNU065-190301-403-d07 CNU065 Homo sapiens cDNA, mRNA sequence.				

ORGANISM	REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE PUBMED	COMMENT
Homo sapiens Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 398) Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.U.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000) 20202663 10737800	Contact: Simpson A.J.G. Laboratory of Cancer Genetics

FEATURES	SOURCE	COMMENT
1. 398	Location/Qualifiers	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL5&t2=IL5-CN0065-100301-403-d07&t3=2001-03-19&t4=1 Seq primer: puc 18 forward High quality sequence stop: 370.

Query Match	0.8%	Score 40.6	DB 12	Length 380
Best Local Similarity	50.8%	Pred. No. 3.4		
Matches 97	Conservative 0	Mismatches 94	Indels 0	Gaps 0
QY	1474	GATGTGAGCATCTCTCGTGGCTTACAGCAAGAAAGTGAAGAGTGGCTGCACGGGAAG	1533	

Query Match	0.8%	Score 40.6	DB 12	Length 453
Best Local Similarity	50.8%	Pred. No. 3.9		
Matches	97	Mismatches	94	Indels
				Gaps
07	1474	GATGTGAGCATCTCGCGCTTACAGAAACCTGAAGAATGGCTCGACCGGAAG	1533	

Db	103	GAGGAGGAGGAGGAGGAATTTCAGCACAGCTTTGAAGAAAGAAAGAAAGAAAGA	162
QY	1534	ATCAGAGAAGCCCTCCACCCCTGCTCCCTGAAATAGAAAAAGAACCGTAGAGCCAGA	1593
Db	163	GAAAGCTGTATGTCGAAGCAAGACTCCAGCAGAAAGAAAGAAAGAAAGCTGTAGTCGAAACA	222
QY	1594	GTAGACCTCATTTATGCAAGAGCCAGAGCAGGTACCTGTGAGACACACAGAGAGACATC	1653
Db	223	GAAAGTCGACGAAGAAAGAAAGAAAGACACAGGAGAGGAGACCGCATAGAAAGAAAGAGA	282
QY	1654	AGGGTGACAAG	1664
Db	283	AGAAAGAGAG	293

RESULT 11

LOCUS	Bj198768	455 bp	mRNA	linear	EST 24-JAN-2002
DEFINITION	Bj198768 normalized full length cDNA library, chloronemata, calionemata and rhizoid-like procnemata Physcomitrella patens subsp. patens cDNA clone pphm22c20 5', mRNA sequence.				
ACCESSION	Bj198768				

ACCESSION	BJ198768
VERSION	BJ198768.1
KEYWORDS	GI:18366689 EST.

SOURCE ORGANISM	physcomitrella patens subsp. patens	physcomitrella patens subsp. patens

REFERENCE
1 (bases 1 to 455)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

AUTHORS
Fujita, T., Shin-1, T., Seki, M., Kamiya, A., Uchiyama, I., Nishiyama, T., Carninci, P., Hayashizaki, Y., Shinzaki, K., Kohara, Y., and Hasebe, M.

TITLE	Comparison of the moss <i>Physcomitrella patens</i> genome with flowering
JOURNAL	plants genome
COMMENT	unpublished
	Contact: Tadaasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

Email: tsahn@genes.nig.ac.jp
A backbone of the vector is pBluescript II, that was in vivo excised from a modified lps phage vector (Mo bi Tac, Germany), XhoI digested-5' end of cDNA is ligated to SalI site of the vector, and the BamHI digested-3' end including poly-A tail is ligated to BamHI site of the vector. cDNA insert could be amplified with conventional T7 and T3 primers. This normalized full-length cDNA library was generated basically according to the method described in Genome Research 10, 1617-1630 (2000).¹ Carnicini, P. et al. Protonemata were blended by the POLYTRON, and then cultivated on the BCD medium containing 1mM NAA (naphthalene acetic acid) for 8 to 11 days under the continuous light.

FEATURES

SOURCE

Source

```

/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pDpn33C20"
/tissue_type="mixture of chloronemata, caulonemata and
rhizoid-like protonemata"
/clone_lib="normalized full length cDNA library"
/chloronemata, caulonemata and rhizoid-like protonemata"
BASE COUNT      138 a      134 g
ORIGIN           111 c      72 t

```

Query Match	0.8%	Score 40.6;	DB 12;	Length 455;
-------------	------	-------------	--------	-------------

Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1474 GATGTGAGCATCTGCGTGCTTACAGCAAGAGCTGAAGAAGTGGCTGCAGCGGAGAG 1533

Db	103	GAGGAGGAGGAGGAGCAAGATTTCAGCCACAGCTTTGAAGAAAGAAAGAAAGAAAGA	162
QY	1514	ATCAGAGAGACCCCTGCCACCTCTGCTCCCTGAAATAGAAAAAGAGCCGTAGAGCAGAA	1593
Db	163	GAAGCTGTAGTCGAAGCAGAGAGTGCAGCAGAAAGAAAGAAAGAAAGCTGTAGTCAGACA	222
QY	1594	GTAGACCTCATTAATGCCAAGAGGCGAGAGCAGTAGCTGTGAGACACACCGAGGACATC	1655
Db	223	GAAGTCGACGAAAGAAAGAAAGAGCAGAGAGGAGCGCATTAAGAAAGAAAGAGA	282
QY	1654	AGCGTGACAG	1664
Db	283	AGAAAGAGAG	293

RESULT 12

LOCUS B2703653 456 bp. mRNA linear EST 25-JAN-2002
DEFINITION B2703653 normalized full length cDNA library, Chloronemeta,
calionemeta and rhinoid-like procnemeta Physcometrella patens
subsp. patens cDNA clone pphn19314 5', mRNA sequence.

ACCESSION	BJ203653	
VERSION	BJ203653.1	GI:18372051
KEYWORDS	EST.	

SOURCE ORGANISM	Physcomitrella patens subsp. patens	Physcomitrella patens subsp. patens
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

REFERENCE
1 (bases 1 to 456)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

AUTHORS
Fujita, T., Shin-i, T., Seki, M., Kamiya, A., Uchiyama, I.,
Carinci, P., Hayaishizaki, Y., Shinozaki, K., Kohara, Y.
, M., and Hasebe

TITLE	Comparison of the moss <i>Physcomitrella patens</i> genome with flowering plants genome
JOURNAL	unpublished
COMMENT	Contact : Tadasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima,
Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

Email: tehini@genes.nig.ac.jp
A backbone of the vector is pBluescript II, that was in vivo excised from a modified lPS phage vector (Mo bi Tec, Germany). XbaI digested-5' end of cDNA is ligated to SalI site of the vector, and the BamHI digested-3' end including poly-A tail is ligated to BamHI site of the vector. cDNA insert could be amplified with conventional T7 and T3 primers. This normalized full-length cDNA library was generated basically according to the method described in Genome Research 10, 1617-1630 (2000). Carninci, P. et al. Protonemata were blended by the POLYTRON, and then cultivated on the BCD medium containing 1mM NaHCO₃ (naphthalene acetic acid) for 8 to 11 days under the continuous light.

FEATURES
source

Source

```

/organism="Phycomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pphn43j14"
/tissue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"
/clone_1lb="normalized full length cDNA library",
chloronemata, caulonemata and rhizoid-like protonemata"
BASE COUNT      138 a      134 g      73 c
ORIGIN

```

Query Match	0.88; Score 40.6; DB 12; Length 456;
-------------	--------------------------------------

Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1474 GATGTGAGCATCTGCGTGCTTACAGCAAGAGCTGAAGAGTGGCTGCAGCGAAGAG 1533

Db	103	GAGGAGGAGGAGGAGCAAGATTTCAGACACAGCTTGAAGAAAGAAAGAAAGAAAGA	162
Qy	1534	ATCAGAGAAAGCCCTGCCACCCCTTCTCCTTAAATATGAAAAAGAGACCGGTAGAGCAGAA	1593
Db	163	GAAGCTGTATGTGGAAGCAGAAATTCGACGAAAGAAAGAAAGAAAGCTGTAGTCGAAGCA	222
Qy	1594	GTAGACCTCATTTATGACAGAGCAGAGCAGAGTAGCTGTGAGACACACACAGAGCACATC	1653
Db	223	GAAGTCCAGCAAGAAAGAAAGAAAGAACACACAGAGAGAGAGACGCTATAGAAAGAAAGA	282
Qy	1654	AGGCTGACAG 1664	
Db	283	AGAAAGAGAG 293	
RESULT 13			
BI195766			
LOCUS	BI195766	715 bp	mRNA linear EST 10-JUL-2001
DEFINITION	602754145F1 NIH_MGC_19 Homo sapiens CDNA clone IMAGE:4889938 5',		
ACCESSION	BI195766		
VERSION	BI195766.1	GI:14650786	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 715)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILML at: Image.1@nl.gov Plate: ILML166 row: 1 column: 11 High quality sequence stop: 556. Location/Qualifiers		
FEATURES			
source	1. .715		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4889938"		
	/tissue_type="neuroblastoma"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH MGC 19"		
	/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."		
BASE COUNT	202 a 194 c 251 g 68 t		
ORIGIN			
Query Match	0.8%; Score 40.4; DB 12; Length 715;		
Best Local Similarity	50.5%; Pred. No. 6.5;		
Matches	98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;		
Qy	1445	TCAAACCCGACCCGGCTATTCAATATGCCGATGTGAGACATCTGCGTGGCTTACGCAAG	1504
Db	223	TTTAAACGACGACGACAGCTGACGACGAGCAGCAGACCCGAGGACACATCAA	282
Qy	1505	AAGCTGAAGAGTGGCTGACGCGGAGAGATCAGAGAACCCCTGCAACCTTTCCTCCTG	1564
Db	283	CACCTGCTCACACACGGGAGGGGGCGCATTCAGAGGCAAGAAAGAGAGGCGCGCGCTG	342

OY	1565	AAATAGAAAAAAGAACCGGTAGAGCCGACGAAGAAGTACCTCATATATGCAGAGCAGACGAG	1622
Dd	343	GAGAGCAACAACGGCCGGAGCCGAGCAGCAGCGGAGACTGTCAGAGAAAGAGCACACGCCG	402
OY	1625	GTACCGTGAGACACA	1638
Dd	403	CGCGCTGGAGAGACA	416
RESULT 14			
LOCUS	BX376097/c	1201 bp	mRNA linear EST 08-MAY-2003
DEFINITION	BX376097 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens		
VERSION	CDNA clone CS0DC022YM12 5'-PRIME, mRNA sequence.		
KEYWORDS	BX376097.1 GI:30434756		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1201) Li,W.B., Gruber,C., Jeesee,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seque@genoscope.cns.fr, web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2866.f http://fulllength.invitrogen.com/InvitrogenCorporation.1600 Faraday Avenue Genoscope sequence ID : CS0DC022B606071.		
FEATURES			
source	Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DC022YM12" /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED" /clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-cligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		
BASE COUNT	79 a 152 c 66 g 270 t 634 others		
ORIGIN			
Query Match	0.8%; Score 40.4; DB 13; Length 1201;		
Best Local Similarity	13.7%; Pred. No. 9.7;		
Matches	135; Conservative 297; Mismatches 551; Indels 2; Gaps 2;		
OY	1487	TGCGTGCCTTACAGCAAGAGCTGAAGAGTGCCTGACGCCGAGAGATCAGAGAACCC	1546
Dd	1186	KRKGGKGDPAADAAAAAAAAAAGKAAAADAADAKMGKKAAGAGAAAKKARAKAGK	11273
OY	1547	TGCCACCCTTCTCCCTGGAATGAAAAAGACCGTAGGCGCAGAGTAGTACCTCATTA	1606
Dd	1126	ARGAAAAAGKAGAKGAGAGGAGGARAKARAGATAWGTAAAGKKRKAAPAAKKRAKDAD	10673
OY	1607	TGCAGAAGGCGACGAGAGCTAGCTGTGAGACACACAGCAGACATCAGGCTGCAAGTT	1666
Dd	1066	GGKAAKKGKMAAGAAAGKAGGKGGKGAABAAKKTMTMMNNNNNNNNNNNNNNNNNN	10077
OY	1667	ACCAAGCGAAGAGAGATTGGGTCTTACGCTATCTTACACCCAGCGGTATTGAATA	1786
Dd	1006	MMAAMMMMMMMHKHHMKKKKHHMHHTMTVTTKMKTKTMMNNNNNNNNNNNNNNNN	947
OY	1727	GTGAAAAAGTGGGTGTATCCACCATTTGGCGGGAACAAGTCTGTATATGCTCACAAAG	1786
Dd	946	KMAAMMMNYMMMTGMNMMNGKKNMMNVMMNNNNNNNNNNNNNNNNNNNNNNNNNN	887

```

Qy 1787 GTAGGCGAGGAGATCAAAATCGACCATACACGGTAAGTCAATTGTACCAGAAAGGA 1846
Db 886 GMMVWYGRGGVMTMGTKTTGKNNMMKMMGMMGNTDMWKTMTMMKMMHMTMTGN 827
Qy 1847 CGGCGGTCCCTGTTCAACATTCGAGCATTTGAGAGCGGCTACGATCTTTTCAACG 1906
Db 826 GKTWNGKNNMMVKKKTTMMVGAATAADGGGAAAAAAGAGAGAGAGAGAGAGAGAG 767
Qy 1907 AGAGGAGATTCGTAACAGATTAACGACATGACATGACATGACGAGAGAGAGAGAGAG 1966
Db 766 NMKMTMTMTMMGTTGKNNMMKMMGNTMTMMKMMGMMGNTMTMMKMMHMTMTGN 707
Qy 1967 CTGACGAAAGTACTATTAAGACTGTAAAGAC-TCAGAGACAGACTGAAATACGCTTTC 2025
Db 706 MKMTMTMTGTTMMTMTMMGTTGKNNMMKMMGNTMTMMKMMGNTMTMMKMMHMTMTGN 647
Qy 2026 GATATTCGACGACGAAAGTGTGTTAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 2085
Db 646 NMKMTMTMTMMGTTGKNNMMKMMGNTMTMMKMMGNTMTMMKMMHMTMTGN 587
Qy 2086 GATCTGTAGATTCACCATTTTCAGAGTTTGCGTACGAGAGAGAGAGAGAGAGAGAG 2145
Db 586 NMKMTMTMTMMGTTGKNNMMKMMGNTMTMMKMMGNTMTMMKMMHMTMTGN 528
Qy 2146 GCACCTCACAAAGTCCCAACATCGAGCTGTAGAGAGAGAGAGAGAGAGAGAGAGAG 2205
Db 527 NMKMTMTMTMMGTTGKNNMMKMMGNTMTMMKMMGNTMTMMKMMHMTMTGN 468
Qy 2206 ATCATCAAAAGCGCTGTACTTAAAGATCTGTTGAGTGCAGAGAGAGAGAGAGAG 2265
Db 467 GNGMMNNMMNNMMNNMMNNMMNNMMNNMMNNMMNNMMNNMMNNMMNNMMNNMMNN 408
Qy 2266 GCAGAAATCATCAGGAGATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2325
Db 407 NMKMTMTMTMMGTTGKNNMMKMMGNTMTMMKMMGNTMTMMKMMHMTMTGN 348
Qy 2326 GATTCAGTCTCTTAATGAGGTTAAGCACCCTGTTAACTCTGTATCATGATGAGCA 2385
Db 347 NAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 288
Qy 2386 TTTCCTGCTCCATGACGAGAGCGCTGTGCTGCTGATGCTGCTGCTGCTGCTGCTG 2445
Db 287 GKAKNNKKAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 228
Qy 2446 GTATTGTGCGGGAACCAAAACAAT 2470
Db 227 GTATTAAGAACCCACCAAAACAAT 203

RESULT 15
BE017217 634 bp mRNA linear EST 30-AUG-2000
LOCUS FK78b02.y1 Zebrafish Research Genetics C32 fin Danio rerio cDNA 5'
DEFINITION similar to SW:ABE_BRARE 042364 APOLOPROTEIN E PRECURSOR.; mRNA
sequence.
ACCESSION BE017217
VERSION BE017217.1 GI:8281665
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 634)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
TITLE Wabhu Zebrafish EST Project 1998
JOURNAL Unpublished
COMMENT Contact: Stephen L. Johnson

```

```

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbratish@wustl.edu
cDNA Library Preparation: Ning Wu. cDNA Library Arrayed by:
Research Genetics. DNA Sequencing by: Washington University Genome
Sequencing Center Clone Distribution: Research Genetics web
address: http://www.researchgenetics.com/
Seq primer: T3 RT from Amersham
High quality sequence stop: 1.
location/Qualifiers
1..634
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/issue_type="Fin"
/lab_host="GeneHogs (HS996, a phage-resistant isolate of
DH10B)"
/clone_lib="Zebrafish Research Genetics C32 fin"
/notes="Vector: pTZ19-Pac with a modified polylinker;
Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was prepared
from zebrafish(C32) fin, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pTZ19 vector. Library is non-normalized. Library was
constructed by Ning Wu. NOTE: This clone is available
royalty-free through LNL; contact the IMAGE Consortium
(info.lnl.gov) for further information"
BASE COUNT 190 a 170 c 142 g 132 t
ORIGIN
Query Match 0.84; Score 40.2; DB 10; Length 634;
Best Local Similarity 50.3%; Pred. No. 6.7;
Matches 99; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
Qy 4115 TGTACCGAATGCGCGGCTGTTGATGACAGCCCAATGCTGTGCGGAGCGGTAGAC 4174
Db 60 TGGACCGATTATGCGCGGCTGTTGATGACAGCCCAATGCTGTGCGGAGCGGTAGAC 119
Qy 4175 TTGTGAAGCAGAACCGCTCATCATGCTGTGAGACCCCAATTTTCTAAGATGCGCG 4234
Db 120 ACATCAAGGCGCTCCCACTCATCATGAGCTTGAAGCACTAATTAAGACATGCTG 179
Qy 4235 AACCGAGGCGGACCTTAAGCTCGACGCTGCTAATGAGCATAGCGTCAACG 4294
Db 180 AACTGAGCTCATACAGTGAATAATATCAAAACCCACATGACCTCATCTTGTATGCTG 239
Qy 4295 CTGAGCGGATTACAAA 4311
Db 240 CTGGTCAGATCAGAAA 256

```

Search completed: November 15, 2003, 19:56:26
Job time : 9001 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 22:31:43 ; Search time 12073 Seconds
(without alignments)
13055.125 Million cell updates/sec

Title: US-10-023-649-1_COPY_5000_11484

Perfect score: 6485

Sequence: 1 caccgcgcgcgcgcgcacac.....catttgccttaaacattc 6485

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_estba:*
3: em_estba:*
4: em_estba:*
5: em_estba:*
6: em_estba:*
7: em_estba:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estfun:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match length	ID	Description
1	58	0.9	985 13	BX436838 BX436838
2	53.6	0.8	836 29	AL302021 Tetrarodon
3	53.4	0.8	1101 29	CNS003DQ
4	52	0.8	1146 29	CNS021G2

Result No.	Score	Match length	ID	Description
5	50.8	0.8	1201 13	BX358198
6	50.4	0.8	997 29	CNS005TE
7	50.2	0.8	1101 29	CNS00182P
8	49.4	0.8	460 14	CB472367
9	49	0.8	844 29	AG042177
10	49	0.8	1101 29	CNS00E7Y
11	48.8	0.8	1165 13	BX338369
12	48.8	0.8	1198 13	BX395612
13	48.8	0.8	1201 13	BX385348
14	48.2	0.7	440 13	BU497603
15	48.2	0.7	566 12	BM275442
16	48.2	0.7	569 12	BM273626
17	48	0.7	887 13	BX441520
18	47.8	0.7	946 13	BX459436
19	47.6	0.7	993 13	B0729165
20	47.4	0.7	1065 9	AL515373
21	47.4	0.7	1101 29	CNS0021U
22	47.4	0.7	1225 29	CNS0161D
23	47.2	0.7	240 9	AU074526
24	47.2	0.7	944 29	CNS06WVA
25	47.2	0.7	994 13	BX414650
26	47.2	0.7	1005 13	BX414452
27	47.2	0.7	1201 9	AL548154
28	47	0.7	712 13	BX416727
29	47	0.7	1201 9	AL558113
30	46.8	0.7	905 29	CNS00KHX
31	46.8	0.7	913 28	AZ546811
32	46.8	0.7	975 13	BX419927
33	46.8	0.7	994 29	CNS015XG
34	46.6	0.7	589 28	AQ541309
35	46.6	0.7	794 28	BH498125
36	46.6	0.7	964 29	CNS017VX
37	46.6	0.7	1101 29	CNS0183Y
38	46.6	0.7	1161 13	BX332029
39	46.6	0.7	1201 13	BX39692
40	46.4	0.7	923 28	AZ538694
41	46.4	0.7	1101 29	CNS003DX
42	46.4	0.7	1201 13	BX316719
43	46.4	0.7	1201 29	CNS016AD
44	46.2	0.7	1101 29	CNS003DQ
45	46.2	0.7	1167 29	CNS07360

ALIGNMENTS

RESULT 1
BX436838
LOCUS
DEFINITION BX436838 Homo sapiens THYMUS Homo sapiens CDNA clone CSOCAP003YH18
3-PRIME, mRNA sequence.
ACCESSION BX436838
VERSION BX436838
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifeotech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CSOCAP003D09NP1.
location/Qualifiers
1. .985
/organism="Homo sapiens"

	BASE COUNT	337 a	120 c	95 g	277 t	156 others
ORIGIN						
	Query Match	0.9%	Score 58;	DB 13;	Length 985;	
	Best Local Similarity	46.7%;	Pred. No. 5.5;			
	Matches	84;	Conservative	27;	Mismatches 69;	Indels 0; Gaps 0;
		Library was not normalized."				
Oy	6303	TAAAGCTGTGACAGTGCATCAATTAAACGATTAATTCACATCACAATCATATTAACACATATACAC 63622				
Dd	746	WMAVATTTAAAAAATAATATATATATATATATATATATAMTTWTATTTTTTMMWATMTAT 805				
Oy	6363	CTTTATGAGACTGACTATGGGTCTATATATACATCACATATATTTTACTTAAAACAC 64222				
Dd	806	ATTTTAAATTAATMMWATTTATATATATATAMAATTAATAATTTATATAMATATAMAATATAT 865				
Oy	6423	TAAACACCTTTATDAATCTTTTATATAATTTTCTTTTGTTTATTTTGTGTTTAAAT 64822				
Dd	866	ATTWTATATATATAAAAAATAATATATATATTTTATATMHTTTWTATATATHTTTTAAAAAT 925				
RESULT 2	CNS04Q18/c					
LOCUS	CNS04Q18	836 bp	DNA	linear	GSS 01-SEP-2000	
DEFINITION	Tetracodon nigroviridis genome survey sequence PUC-Ori end of clone					
	128L13 of library G from Tetracodon nigroviridis, genomic survey					
	sequence.					
ACCESSION	AL302021					
VERSION	AL302021.1	GI:8179132				
KEYWORDS	GSS; genome survey sequence.					
SOURCE	Tetracodon nigroviridis					
ORGANISM	Tetracodon nigroviridis					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;					
	Acanthomorphi; Acanthopterygii; Perccomphi; Tetraodontiformes;					
	Tetraodontidae; Tetraodontidae; Tetraodon.					
REFERENCE	1					
AUTHORS	Roeest Crollius,H., Jailion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Winkler,P., Brotier,P., Quetier,F., Saurin,W. and Weissenbach,J.					
TITLE	Estimate of human gene number provided by genome-wide analysis					
JOURNAL	using Tetracodon nigroviridis DNA sequence					
MEDLINE	Nat. Genet. 25 (2), 235-238 (2000)					
PUBMED	20296633					
REFERENCE	10835645					
AUTHORS	2					
	Roeest Crollius,H., Jailion,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.					
TITLE	Characterization and repeat analysis of the compact genome of the					
JOURNAL	freshwater pultefish Tetracodon nigroviridis					
MEDLINE	Genome Res. 10 (7), 939-949 (2000)					
PUBMED	10899143					
REFERENCE	3	(bases 1 to 836)				
AUTHORS	Genoscope.					
TITLE	Direct Submission					
JOURNAL	Submitted (12-Apr-2000) Genoscope - Centre National de Sequencage :					
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr					
	Web : www.genoscope.cns.fr)					
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetracodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetracodon.					

FEATURES	source	1. .836	location/Qualifiers	
			/organism="Tetraneura nigroviridis"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:99883"	
			/clone="128L13"	
			/clone_1lb="G"	
			/note="Genoscope sequence ID : COBG128CF07SP1-end ; PUC-Or1"	
BASE COUNT	301 a	30 c	173 g	299 t 33 others
ORIGIN				
Query Match	0.8%	Score 53.6	DB 29	Length 836
Best Local Similarity	57.8%	Pred. No. 33		
Matches	89	Conservative	3	Mismatches 62; Indels 0; Gaps 0;
Qy	6331	ATATCACACTACATATTTAACACACTATATTCATCTTTATGAGACTCACTATAGGCTCTCTA	6330	
Db	579	ATTATTAATCTTAATTAATAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	520	
Qy	6391	ATATACACTACATATTTTACTTAATAAACCTATACACTTTATTAATTCCTTTATTA	6450	
Db	519	ATATCAATTTTACATTAATTAATTAATTAATAAATTAATTAATTAATTAATTAATTAAT	460	
Qy	6451	TTTTCTTTGTTTATTTATTTGTTTAAATTT	6484	
Db	459	TTAATTTATTTATTTATTTATTTTTCATATTT	426	
RESULT 3				
CNS003DQ				
LOCUS	CNS003DQ	1101 bp	DNA	linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #			
	BACR08109 of RPCI-98 library from Drosophila melanogaster (fruit			
	fly), genomic survey sequence.			
ACCESSION	AL064580			
VERSION	AL064580.1	GI:4941932		
KEYWORDS				
SOURCE				
ORGANISM	Drosophila melanogaster (fruit fly)			
	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
	Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 1101)			
AUTHORS	Genoscope.			
FILE	Direct Submission			
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segr@genoscope.cns.fr			
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .			
FEATURES	location/Qualifiers			
source	1. .1101			
	/organism="Drosophila melanogaster"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:7227"			
	/clone="BACR08109"			
	/clone_1lb="RPCI-98"			
	/note="end : TET3"			
BASE COUNT	291 a	51 c	117 g	404 t 238 others

ORIGIN

Query Match 0.8%; Score 53.4; DB 29; Length 1101;
Best Local Similarity 51.5%; Pred. No. 33;
Matches 84; Conservative 14; Mismatches 65; Indels 0; Gaps 0;

QY 6332 ATTAACTGATATACACTACATATTAAACAACATATACCTTTATGAGCTGACTAT 6381
Db 274 ATWATWTATATWMAWMAWTAATATAAAMAAAAATWMAAATATATAAATAATATMT 333
QY 6382 GGGTCTCAATATACATACATACATATTCTTAAACACATATACACTTTAATAT 6441
Db 334 WMAATWTWTAATATATATATATATATATATATATATATATATATATATATATAT 393
QY 6442 CTTTATATATATTTCTTTGTTTATTTGTTTATTTAAATTT 6484
Db 394 ATTTTATATATTTTATTTTATTTTATTTTATTTATTTATTTMT 436

RESULT 4
CNS021G2 1146 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 225004 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL176843.1 GI:7814900
VERSION AL176843
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphi; Acanthopterygii; Percormorphi; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1
AUTHORS Roest Croollius, H., Jallou, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fitzames, C., Wincker, P., Brothier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
ESTIMATE of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

TITLE JOURNAL
MEDLINE 20296633
PUBMED 10835645

REFERENCE 2
AUTHORS Roest Croollius, H., Jallou, O., Dasilva, C., Ozouf-Costaz, C.,
Fitzames, C., Fischer, C., Bouneau, L., Billaud, A., Quetier, F.,
Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

JOURNAL 20359837
MEDLINE 10899143
PUBMED 3 (bases 1 to 1146)
REFERENCE Genoscope.
AUTHORS Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-and-sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
SOURCE location/Qualifiers
1. 1146
/organism="Tetraodon nigroviridis"
/mol_type="Genomic DNA"
/db_xref="taxon:99883"
/clone="225004"
/clone_lib="G"
/note="Genoscope sequence ID : CNS0225BH02LPI-end : T7"

BASE COUNT 393 a 60 c 73 g 274 t 346 others
ORIGIN

Query Match 0.8%; Score 52; DB 29; Length 1146;
Best Local Similarity 31.7%; Pred. No. 57;
Matches 113; Conservative 71; Mismatches 172; Indels 0; Gaps 0;

QY 6126 TTGTTGAGCACTTATAGTGTGCTGACGCTTATGCTTATTAACACAGTATGAC 6185
Db 692 TTDTTTAAKKKKKKKKKKAGGKKRKRKRKRKRKRKRKRKRKRKRKRKRKRKRKRKR 751
QY 6186 TGAGCGCGGACACGACGACGACGACGACGACGACGACGACGACGACGACGACG 6245
Db 752 AAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 811
QY 6246 AATGCAACGCCCGCTGACACTAACTCGATGATTTCCGAGAGACAGTGCATAA 6305
Db 812 GRRKKKKGGDMATTTTDTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 871
QY 6306 TCGTGTGACGTGCACATTAATCGTATATCACTACATATTAAACAACATATACCTT 6365
Db 872 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 931
QY 6366 TTATGAGACGACATGCGCTCTAATATACATACATATTTTACTTAAACACATAT 6425
Db 932 TTAATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 991
QY 6426 ACACACTTATATATTTCTTTATATATTTCTTTGTTTATTTGTTTATTTTAA 6481
Db 992 TATTTTATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1047

RESULT 5
BX358198/c 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX358198 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CSOD1035AC19 5-PRIME, mRNA sequence.
ACCESSION BX358198.1 GI:30368196
VERSION BX358198
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL Contact: Genoscope
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1590.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1035AC10P1&cluster=1590.f. Contact :
Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue genoscope sequence ID : CSOD1035AC10P1.

TITLE JOURNAL
MEDLINE 11201
PUBMED 1 (bases 1 to 1201)
REFERENCE Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1035AC19"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand CDNA was primed with a NotI-oligo (dt)
primer. Five prime end enriched, double-strand CDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 373 a 144 c 192 g 367 t 125 others
ORIGIN

Query Match 0.8%; Score 50.8; DB 13; Length 1201;

Db 196 ATTTTTCCTTAATTTAAATTTTTTTTCCTTTTAAAGTTTTTTTACCTTTTTTGGG 137

Dy 6374 CTCACATGGGCTCTAATATATACACATATTTTACTTTAAACACACTATACACATT 6433

D_b **136 TATTTTTTTTTTTT TTTT TTTT TTTT TTA AAAA CTTTTTTTGTTT 77**

6434 TATAAATCTTTATAAATTTCTTTGTTTATAATTTGTTTAAAAATTT 6484

AG042177/c			
LOCUS	844 bp	DNA	linear
DEFINITION	Pan troglodytes DNA, clone: PB-02B19.F, genomic survey sequence.	GSS	01-NOV-2000

```

VERSION      AG042177.1  GI:16570902
KEYWORDS
SOURCE       GSS.
             Pan troglodytes (chimpanzee)

```

REFERENCE

1 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

Totoki, Y., Watanabe, H. and Sakaki, Y.

JOURNAL
Unpublished
0 4-1111 1 4-1 111

AUTHORS
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki Y. Watanabe H. and Sakaki Y.

TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9111, Fax: 81-45-503-9170

was generated during the R&D process and may have higher chance of

PRIMERS
Sequencing: -31M13

LIBRARY
Vector : PKS145

```
R:Site 1 :: Sact
R:Site 2 :: Sact.
```

FEATURES	LOCATION/QUALIFIERS
SOURCE	1. .844

```

/organism="H. sapiens"
/mol_type="genomic DNA"

```

```

/clone="PTB-020B19.F"

```

```

/cell_type="lymphoblast"
/cldonlib="PDB Chimera200 Majo BAC 1ibaw"

```

BASE COUNT	243 a	43 c	66 g	487 t	5 others	.
ORIGIN						

Query Match 0.84: Score 49: DB 29: Length 844:

```

Best Local Similarity  54.4%;  Freq. NO. 26702;
Matches  112;  Conservative  0;  Mismatches  106;  Indels  0;  Gaps  0;

```

6264 A C A C T A A A C T G A T G T A T T C C G A G G A A C C A C A G T G C A T A A T G C T G T C A G T G T C A C A T 6323

Db 601 ATAAGAAATTATATATATAATAAGCAATTATACAAATAAAAAAAAAAGTNTATT 542

QY 6324 TAATCGTATATCACACTACATATTTAACCAACACTATATCACTTTTATGAGACTCACTATGG 6383

D5 541 AAATAATAAAAAATATAATATAAAAAAATAATAATTTATTAATAAAAAATATAA 482

[illegible]

.....

OY	6444	TTTAAATTTCTTTGTTTATTGTTTAAAA	6481
Db	421	TTAAAATAAGTATTTTTTAATTTTTTAAATAA	384
RESULT 10			
CNS00E7Y		1101 bp	DNA
LOCUS			linear GSS 04-JUN-1999
DEFINITION			Drosophila melanogaster genome survey sequence TET3 end of BAC #
ACCESSION			BACR28A16 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION			A1069057
KEYWORDS			AL069057.1 GI:4949017
SOURCE			GSS.
ORGANISM			Drosophila melanogaster (fruit fly)
			Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridae; Drosophilidae; Drosophila.
REFERENCE			1 (bases 1 to 1101)
AUTHORS			Genoscope.
TITLE			Direct Submission
JOURNAL			Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT			Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mammossier in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .
FEATURES			Location/Qualifiers
source			1..1101
			/organism="Drosophila melanogaster"
			/mol_type="genomic DNA"
			/db_xref="taxon:7227"
			/clone="BACR28A16"
			/clone_id="RPCL-98"
			/note="end : TET3"
BASE COUNT			248 a 175 c 173 g 341 t 164 others
ORIGIN			
Query Match			0.8%; Score 49; DB 29; Length 1101;
Bacc Local Similarity			38.1%; Fred. No. 1.9e+02;
Matches			61; Conservative 44; Mismatches 55; Indels 0; Gaps 0;
OY	6318	TCACATTATGCTATGCACACTACATATTAAACAACATATCACTTTTAGAGACCA	6377
Db	910	TCCAAAAHACHMHAHTCHTMWMTATTTTMAATTTTWTMTTTTTTTTAAATTTWA	969
OY	6378	CTATGGGTCTGTATATACACTACACATATTTTACTTAAAAAACATATACACTTATA	6437
Db	970	AMHTTHMMTMAAAAATATTTTATTAHAAATMTMTTTTMMAMWMTAMWHHTCWNY	1029
OY	6438	AATCTTTTATATTTTCTTTTGTTTATTTTGTTTT 6477	
Db	1030	CMTTTTTTTWTMTMTHTMTMTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1069	
RESULT 11			
EXJ38369/c		1165 bp	mRNA linear EST 02-MAY-2003
LOCUS			EXJ38369 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA DEFINITION

ACCESSION	clone CSODI058YF14 5-PRIME, mRNA sequence.
VERSION	BX338369
KEYWORDS	BX338369.1 GI:30343682
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	L.I.W.B., Gruber.C., Jeesee.J. and Polayes.D.
TITLE	(bases 1 to 1165)
JOURNAL	Full-length cDNA libraries and normalization unpublished
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODI058DC070P1. Location/Qualifiers 1..1165
FEATURES	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODI058YF14" /cissue_type="PLACENTA COT 25-NORMALIZED" /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-clig0 (dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT	567 a 109 c 140 g 251 t 98 others
ORIGIN	
Query Match	0.8%; Score 49; DB 13; Length 1165;
Best Local Similarity	41.1%; Pred. No. 1.8e+02;
Matches	67; Conservative 33; Mismatches 63; Indels 0; Gaps 0;
Dy	6322 ATTATCGTATATGACACTCATATTAAACAAGACTATTCCTTTATGAGACTCACTAT 6381 ::: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: 977 WMTWTWMAAAMAAWATAATWMAATWMAATWMAATWMAATWMAATWMAATWMAATWMAAT 918
Dd	6382 GGCTCTGAATATACACTACACATATTTTACTTAAAAAGACTATACACCTTTTAATT 6441 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: 917 WTTTATATATTTTWTATTAATTTWTTTATATWTAATAATATWTAAMAAMAAAATAAAT 858
Oy	6442 CTTTATATAATTTCTTTGGTTTTATTTGTTTTAAATTT 6484 ::: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: 857 YWTTWMAATWMAATWMAATWMAATWTTTITTTTITTTTITTTT 815
RESULT 12	
LOCUS	BX395612 1198 bp mRNA linear EST 13-MAY-2003
DEFINITION	BX395612 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI002YP22 3-PRIME, mRNA sequence.
ACCESSION	BX395612
KEYWORDS	BX395612.1 GI:30616740
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	L.I.W.B., Gruber.C., Jeesee.J. and Polayes.D.
TITLE	(bases 1 to 1198)
JOURNAL	Full-length cDNA libraries and normalization unpublished
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of

mRNA isolation system (Promega, WI) using streptavidin
MagneSphere particles. Directional cDNA libraries were
constructed by oligo d(T) priming of poly(A)+ RNA (5mg)
into EcoRI and XhoI sites of 1 ZapII vector using the Zap
cDNA synthesis kit (Stratagene, CA). The average size of the
cDNA inserts in the library was between 1.0 and 1.5kb.
Clones were mass excised using the ExAssist helper phage
(Stratagene), the phagemids were precipitated with PEG
8000 and extracted with phenol/chloroform. Phagemid DNA
was electroporated into DH10B cells."

BASE COUNT 195 a 22 c 14 g 208 t 1 others
ORIGIN

Query Match 0.7%; Score 48.2; DB 13; Length 440;
Best Local Similarity 55.8%; Pred. No. 3.3e+02;
Matches 92; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 6320 ACATTAACTGATATGCACTACATATTAAACACATATACCTTTATGAGACTGACT 6379
DB 215 ATATATATATATATATATATATATATATATATATATATATATATATATAT 274
QY 6380 ATGGCTCTATATATACATACATATTTTACTTAAACACATACACTTATATAA 6439
DB 275 ATTAATATATATATATATATATATATATATATATATATATATATATATAT 334
QY 6440 TTTCTTATATATTTTCTTTGTTTATTTGTTTAAATTT 6484
DB 335 ATATATATATATTTCTTATATATATATATATATATATATATTTT 379

RESULT 15 566 bp mRNA linear EST 20-DEC-2001
BM275442
LOCUS
DEFINITION
BM275442.1 GI:17968772
EST.
Plasmodium falciparum 3D7
Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
1 (bases 1 to 566)
Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Page, D.,
Matta, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B.,
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I.,
Tsagarelis, V., Richey, J., Madkins, J., Kennedy, S., Levinso, D.,
Waterston, R., Wilson, R., and Sibley, D.
Washu Plasmodium EST Project
Unpublished

TITLE
JOURNAL
COMMENT
Contact: L. David Sibley
Washu Plasmodium EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314.286.1800
Fax: 314.286.1810
Email: est@wustl.edu

Library was constructed by R. Haywood. DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: L. David Sibley
(lsibley@jorim.wustl.edu), Washington University
Seq primer: -40UP from Gibco
High quality sequence stop: 423.

FEATURES

source
1. .566
Location/Qualifiers
/organism="Plasmodium falciparum 3D7"
/mol_type="mRNA"
/db_xref="taxon:36329"
/dev_stage="gametocyte (stage III-V)"
/lab_host="DH10B (Genetog, Invitrogen, Inc.)"
/clone_lib="Plasmodium falciparum 3D7 gametocyte cDNA
library"

/note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:
XhoI; The library was constructed by R. Haywood. cDNAs were
synthesized from gametocyte poly(A)+ RNA by oligo d(T)
priming, size-selected and directionally cloned into the
EcoRI (5' end) to XhoI (3' end) sites of the Uni-ZAP XR
lambda vector (Stratagene). The primary library was mass
excised as phagemid using the ExAssist helper phage
(Stratagene). Clones were mass excised using the ExAssist
helper phage (Stratagene), the phagemids were precipitated
with PEG 8000 and extracted with phenol/chloroform.
Phagemid DNA was electroporated into DH10B cells. Clone
Availability: David Sibley, Washington University."

BASE COUNT 256 a 63 c 194 t
ORIGIN

Query Match 0.7%; Score 48.2; DB 12; Length 566;
Best Local Similarity 57.7%; Pred. No. 3.1e+02;
Matches 86; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 6336 ACATCAATATATACAACTATATGACTTTATGAGACTGATGAGTCTTAATATA 6395
DB 1 AGAATATATATATATATATATATATATATATATATATATATATATATATA 60
QY 6396 CACTACATATATTTTACTTAAACACATATACACTTATATAATCTTTATATTTT 6455
DB 61 TAATATATATATATATATATATATATATATATATATATATATATATATA 120
QY 6456 CTTTGTATATTTTGTATTTTAAATTT 6484
DB 121 TATATATCTCTTTTTCACACTT 149

Search completed: November 16, 2003, 08:59:44
Job time : 12079 secs

GenCore version 5.1.6
(c) 1993 - 2003 Compu

Search time 1496 Seconds
(without alignments)
10923.143 Million cell updates/sec

.....atcatcctagaacgccc 5000

residues

ameters: 4339922

★

redicted by chance to have a score of the result being printed
ical score distribution.

TEST

Description	
Sequence 1, April	-649-1
Sequence 1, April	-258-1
Sequence 14, April	-258-14
Sequence 17, April	-258-17
Sequence 2, April	-258-2
Sequence 3, April	-412-3
Sequence 1, April	-106-1
Sequence 1, April	-246-1
Sequence 4, April	-246-4
Sequence 102, April	-362-102
Sequence 103, April	-362-103
Sequence 101, April	-362-101
Sequence 1, April	-042-1
Sequence 1, April	-5-883-1
Sequence 51, April	-101-51
Sequence 8, April	-258-8

```

17 211.4 4.2 6989 10 US-09-191-258-11 Sequence 11, Appl
18 90.8 1.8 282 12 US-09-190-246-7 Sequence 7, Appl
19 50.6 1.0 170 12 US-10-123-101-90 Sequence 90, Appl
20 49 1.0 327 12 US-10-123-101-65 Sequence 65, Appl
21 49 1.0 346 12 US-10-123-101-67 Sequence 67, Appl
22 48.6 1.0 132 12 US-10-123-101-68 Sequence 68, Appl
23 48.6 0.8 416 12 US-10-027-632-80186 Sequence 80186, A
24 40 0.8 416 12 US-10-027-632-80187 Sequence 80187, A
25 40 0.8 416 12 US-10-027-632-315015 Sequence 315015,
26 40 0.8 416 12 US-10-027-632-315016 Sequence 315016,
27 40 0.8 416 13 US-10-027-632-80186 Sequence 80186, A
28 40 0.8 416 13 US-10-027-632-80187 Sequence 80187, A
29 40 0.8 416 13 US-10-027-632-315015 Sequence 315015,
30 40 0.8 416 13 US-10-027-632-315016 Sequence 315016,
31 39.6 0.8 4610 10 US-09-884-001-1 Sequence 1, Appl
32 39.6 0.8 6479 13 US-10-098-841-101 Sequence 101, App
33 39.6 0.8 775 12 US-10-140-472-120 Sequence 120, App
34 39.2 0.8 775 12 US-10-141-761-120 Sequence 120, App
35 39.2 0.8 775 12 US-10-142-885-120 Sequence 120, App
36 39.2 0.8 775 12 US-10-158-790-120 Sequence 120, App
37 39.2 0.8 775 12 US-10-137-871-120 Sequence 120, App
38 39.2 0.8 775 12 US-10-140-805-120 Sequence 120, App
39 39.2 0.8 775 12 US-10-140-864-120 Sequence 120, App
40 39.2 0.8 775 12 US-10-140-923-120 Sequence 120, App
41 39.2 0.8 775 12 US-10-141-756-120 Sequence 120, App
42 39.2 0.8 775 12 US-10-141-759-120 Sequence 120, App
43 39.2 0.8 775 14 US-10-123-155-120 Sequence 120, App
44 39.2 0.8 775 15 US-10-146-731-120 Sequence 120, App
45 39.2 0.8 1184 12 US-10-140-472-412 Sequence 412, App

ALIGNMENTS

RESULT 1
US-10-023-649-1
Sequence 1, Application US/10023649
Publication No. US20030143201A1
GENERAL INFORMATION:
APPLICANT: The Minister of National Defence, Government of Canada
APPLICANT: Nagata, Leslie P
APPLICANT: Wong, Jonathan P
TITLE OF INVENTION: No. US20030143201A1el DNA-Based Vaccine Against the Encephalitis
FILE REFERENCE: NEI-001
CURRENT APPLICATION NUMBER: US/10/023, 649
CURRENT FILING DATE: 2002-12-27
PRIOR APPLICATION NUMBER: 60/256, 948
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1.
LENGTH: 11484
TYPE: DNA
ORGANISM: Western equine encephalomyelitis virus - strain 71V-1658
FEATURES:
NAME/KEY: CDS
LOCATION: (25)..(7428)
OTHER INFORMATION: 5' UTR <1 .. 24
FEATURE:
NAME/KEY: CDS
LOCATION: (7473)..(11183)
OTHER INFORMATION: encodes nucleocapsid, E3, E2, 6K and E1 proteins
US-10-023-649-1

Query Match 100.0%; Score 5000; DB 12; Length 11484;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ACCCTACAACCTAATGATCCATATGGAAGAATTCAGTTGACTTGATGCTGACACG 60
Db 1 ACCCTACAACCTAATGATCCATATGGAAGAATTCAGTTGACTTGATGCTGACACG 60
Oy 61 CCGATATGCAAGTCGTTACAGCGAGCGTTTCACCAATTGTGATGGAAGCAAGCAGGCTC 120

```

Db	61	CCGATGTCAGTGTATACGCGACGTTTCCAACTTTGAGTGAAGCAAGGCAAGTCTC	120
Qy	121	ACTGACAAATGCAATGCCAATATGCAAGCGTTTTCGATGTGGCAACAAGCTCATTTAG	180
Db	121	ACTGACAAATGCAATGTCCAAATATGCAAGCGTTTTCGATGTGGCAACAAGCTCATTTAG	180
Qy	181	AGCGAAGTCGACCGGGACCAAGTTATCTTGAACATTTGAATGGCCCGCTGACATGTCA	240
Db	181	AGCGAAGTCGACCGGGACCAAGTTATCTTGAACATTTGAATGGCCCGCTGACATGTCA	240
Qy	241	CATTCCAATCACCGCTATCATTTGATCTGCCCTATGATTAAGCGCTGAAGACCCGGACGA	300
Db	241	CATTCCAATCACCGCTATCATTTGATCTGCCCTATGATTAAGCGCTGAAGACCCGGACGA	300
Qy	301	CTACACGCGTATGCAAGAAAGACTTAAGAAAGTGCATTTACCGACAGACATAGCTCT	360
Db	301	CTACACGCGTATGCAAGAAAGACTTAAGAAAGTGCATTTACCGACAGACATAGCTCT	360
Qy	361	AAGGCGGACGCTGCTGGAAGTCACTGTCACACAGAGGAGACTCCATCTCTGTGT	420
Db	361	AAGGCGGACGCTGCTGGAAGTCACTGTCACACAGAGGAGACTCCATCTCTGTGT	420
Qy	421	ATGCAACAGACGCGCAGCTGTAGGTACTTTGAAAGTGTAGCATTAACAGATGTGTAC	480
Db	421	ATGCAACAGACGCGCAGCTGTAGGTACTTTGAAAGTGTAGCATTAACAGATGTGTAC	480
Qy	481	GCAATCCATGCAACGACATCATCTACCAAGCGGCTTAAAGAGTTAGCAATTTAC	540
Db	481	GCAATCCATGCAACGACATCATCTACCAAGCGGCTTAAAGAGTTAGCAATTTAC	540
Qy	541	TGGAATGCGCTTGAACGACCCCTTTATGTACAAAAAATGGGAGGTTCCATCCCTACT	600
Db	541	TGGAATGCGCTTGAACGACCCCTTTATGTACAAAAAATGGGAGGTTCCATCCCTACT	600
Qy	601	TACAAACGAACTGGGCTGACGAGAGAGTATTTGAAAGCAGCATTTGGCTCGGTAC	660
Db	601	TACAAACGAACTGGGCTGACGAGAGAGTATTTGAAAGCAGCATTTGGCTCGGTAC	660
Qy	661	TCAATCTTTCAGAGAGAGGCTTGGAAAATCTTCATCTTTAGAGAGAAGGCTCCAA	720
Db	661	TCAATCTTTCAGAGAGAGGCTTGGAAAATCTTCATCTTTAGAGAGAAGGCTCCAA	720
Qy	721	CTTACTAATAGATCATTTCTCGTTGGTCAACAATCTACACAGAAATAGATCATCTG	780
Db	721	CTTACTAATAGATCATTTCTCGTTGGTCAACAATCTACACAGAAATAGATCATCTG	780
Qy	781	TTAGTAGCTGGCATCTTCCAAAAGTGTTCACATTTGAAGAAAGTCTTACCATG	840
Db	781	TTAGTAGCTGGCATCTTCCAAAAGTGTTCACATTTGAAGAAAGTCTTACCATG	840
Qy	841	AGATGTGGGACCATTTGCTGAGAGGTTACGTCATCAAAAAGATTAAGATCAAGCCCA	900
Db	841	AGATGTGGGACCATTTGCTGAGAGGTTACGTCATCAAAAAGATTAAGATCAAGCCCA	900
Qy	901	AGATGTGGGACCATTTGCTGAGAGGTTACGTCATCAAAAAGATTAAGATCAAGCCCA	900
Db	901	AGATGTGGGACCATTTGCTGAGAGGTTACGTCATCAAAAAGATTAAGATCAAGCCCA	900
Qy	961	TGCAAAAGTCAGATACGCTCGCGGCGAGAGAGGTTCTTTGCTGTGTGATGATGTA	1020
Db	961	TGCAAAAGTCAGATACGCTCGCGGCGAGAGAGGTTCTTTGCTGTGTGATGATGTA	1020
Qy	1021	CCAGCCACACTTTGGCATCAAGATGACAGGATTTCTGGCACTGACGTTTAGTGGATGAC	1080
Db	1021	CCAGCCACACTTTGGCATCAAGATGACAGGATTTCTGGCACTGACGTTTAGTGGATGAC	1080
Qy	1081	GCAAAAAAACTAATGCTTGGGCTCAACCAAGATTTGCTCAATGTGATGATGATGATG	1140
Db	1081	GCAAAAAAACTAATGCTTGGGCTCAACCAAGATTTGCTCAATGTGATGATGATGATG	1140
Qy	1141	AATACTAACAATATGCAAGAACTATCTATTTACAGTGTGCGCCAGGCGTTTCCAGGTG	1200
Db	1141	AATACTAACAATATGCAAGAACTATCTATTTACAGTGTGCGCCAGGCGTTTCCAGGTG	1200
Qy	1201	GCGGTGAACATGCTGCCACTTGGAGACGAGAAAGAACTAAGGGTTCGGGAGCGCACT	1260
Db	1201	GCGGTGAACATGCTGCCACTTGGAGACGAGAAAGAACTAAGGGTTCGGGAGCGCACT	1260
Qy	1261	CTTACTATGGGCTGCTGCTGGGCTTTCAAGACCCAGAAATTCATCATCTACAAAG	1320
Db	1261	CTTACTATGGGCTGCTGCTGGGCTTTCAAGACCCAGAAATTCATCATCTACAAAG	1320
Qy	1321	CCTGTGACGCAACATTAAGAAAGTACCTGCCGCTCTTGAACCTCATTTGTGATTTCAAGC	1380
Db	1321	CCTGTGACGCAACATTAAGAAAGTACCTGCCGCTCTTGAACCTCATTTGTGATTTCAAGC	1380
Qy	1381	CTTACAGCCACGGGCTCGATATGGGCTTCCGCGTGAAGCTCAAGCTGCTTGAACCA	1440
Db	1381	CTTACAGCCACGGGCTCGATATGGGCTTCCGCGTGAAGCTCAAGCTGCTTGAACCA	1440
Qy	1441	ACTGTCAAAACCGGACCGGCTATTACAAATGGCCGATGTGAGCATCTGCGTGGCTTACAG	1500
Db	1441	ACTGTCAAAACCGGACCGGCTATTACAAATGGCCGATGTGAGCATCTGCGTGGCTTACAG	1500
Qy	1501	CAAGAACCTGAAGAGTGGCTGACGCGAAAGATTCAGAGAGCCCTGCAACCTTCTCTC	1560
Db	1501	CAAGAACCTGAAGAGTGGCTGACGCGAAAGATTCAGAGAGCCCTGCAACCTTCTCTC	1560
Qy	1561	CTTGAATTAAGAAAAAGACCTGTATGAGGCAAGATTAACCTCATTTATGCAAGGCGAGA	1620
Db	1561	CTTGAATTAAGAAAAAGACCTGTATGAGGCAAGATTAACCTCATTTATGCAAGGCGAGA	1620
Qy	1621	GCAAGTACGCTGAGACACCAAGAGCAATCATGAGGCTGCAAGTTTCCAGGCGAAGAG	1680
Db	1621	GCAAGTACGCTGAGACACCAAGAGCAATCATGAGGCTGCAAGTTTCCAGGCGAAGAG	1680
Qy	1681	AAGATGGGCTTATGAGCTATCTTTCAACCCAGGCGGATTTGAATGTAAGAAAACTGGCG	1740
Db	1681	AAGATGGGCTTATGAGCTATCTTTCAACCCAGGCGGATTTGAATGTAAGAAAACTGGCG	1740
Qy	1741	TGATTCACCCATTTGGGGAACAAGTACTGTATGACTCAACAAAGGTAGGGCGAGAGA	1800
Db	1741	TGATTCACCCATTTGGGGAACAAGTACTGTATGACTCAACAAAGGTAGGGCGAGAGA	1800
Qy	1801	TACAAATGAGCCATTCACAGGTAAGTCAATTTGTAACAGAAAGGAGCGGCTCTGTT	1860
Db	1801	TACAAATGAGCCATTCACAGGTAAGTCAATTTGTAACAGAAAGGAGCGGCTCTGTT	1860
Qy	1861	CAAGACTTCCAGGCTATGAGTGAAGGCGTACGATCGTTTCAACGAGAGGAGGTTGCTA	1920
Db	1861	CAAGACTTCCAGGCTATGAGTGAAGGCGTACGATCGTTTCAACGAGAGGAGGTTGCTA	1920
Qy	1921	AACGATTAAGCTGCAACATTCGCAATTCACGAGAGGCGCTTAAACCTGACGAAAGTAC	1980
Db	1921	AACGATTAAGCTGCAACATTCGCAATTCACGAGAGGCGCTTAAACCTGACGAAAGTAC	1980
Qy	1981	TATTAAGCTGTAAGACTCAGAGACAGACTCAGAAATCGTCTTCAATTTAGAGCAGCA	2040
Db	1981	TATTAAGCTGTAAGACTCAGAGACAGACTCAGAAATCGTCTTCAATTTAGAGCAGCA	2040
Qy	2041	AAGTGTATTAAGCGAGAAAGCGAGGTCCTTGTGCTTAAACCGGTGATCTGTGATGATCA	2100
Db	2041	AAGTGTATTAAGCGAGAAAGCGAGGTCCTTGTGCTTAAACCGGTGATCTGTGATGATCA	2100
Qy	2101	CCATTTTACAGAGTTTGGCTACAGAGTCTCAAGACAGACCAAGAGACCTTCAAAAGTGC	2160
Db	2101	CCATTTTACAGAGTTTGGCTACAGAGTCTCAAGACAGACCAAGAGACCTTCAAAAGTGC	2160
Qy	2161	CCAAACATCGGAAGCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGGCT	2220
Db	2161	CCAAACATCGGAAGCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGGCT	2220
Qy	2221	GTAAGTAAGAAAGATCTGTGTGTGATGCGAAGAAAGGAAAACTGCGCAGAAATCATCAGG	2280
Db	2221	GTAAGTAAGAAAGATCTGTGTGTGATGCGAAGAAAGGAAAACTGCGCAGAAATCATCAGG	2280

QY 2281 GATGTAAGAGAGATGAGACGTATGATGTTGCTGTAGAGACTGTGATTCAGTGTCTTA 2340
DB 2281 GATGTAAGAGAGATGAGACGTATGATGTTGCTGTAGAGACTGTGATTCAGTGTCTTA 2340
QY 2341 AATGGGGTTAAGACACCCCGTTTAACTCTGTATCATTTATGAGCAATTTGCCCTGATCA 2400
DB 2341 AATGGGGTTAAGACACCCCGTTTAACTCTGTATCATTTATGAGCAATTTGCCCTGATCA 2400
QY 2401 GGGAGCGCTGTGGACCTGATTTGCTTAAAGGATGATTTGTTGGGGGAC 2460
DB 2401 GGGAGCGCTGTGGACCTGATTTGCTTAAAGGATGATTTGTTGGGGGAC 2460
QY 2461 CCAAAACATGCGGCTTCTTAAACATGATGTGCTGAAAGTACATTTTAAACATGACATA 2520
DB 2461 CCAAAACATGCGGCTTCTTAAACATGATGTGCTGAAAGTACATTTTAAACATGACATA 2520
QY 2521 TGCACTGAAGTGTACATTAAGACATCTCTAGAGGTGACACAGACTGTAAACGCCATC 2580
DB 2521 TGCACTGAAGTGTACATTAAGACATCTCTAGAGGTGACACAGACTGTAAACGCCATC 2580
QY 2581 GTCTCCAGCTCTTCTAGACAAAGGATGAAAGCGTTAAACCATGTGCTGATTAATC 2640
DB 2581 GTCTCCAGCTCTTCTAGACAAAGGATGAAAGCGTTAAACCATGTGCTGATTAATC 2640
QY 2641 ATCTAGATATACAGAGGACCAAAAGCGCACAAAGATGATCTGATTTAACTGTTTC 2700
DB 2641 ATCTAGATATACAGAGGACCAAAAGCGCACAAAGATGATCTGATTTAACTGTTTC 2700
QY 2701 AGAGAGATGGTGAAACAGCTACAGATTGACTACAAATAACGAAATCATGCTGCGCT 2760
DB 2701 AGAGAGATGGTGAAACAGCTACAGATTGACTACAAATAACGAAATCATGCTGCGCT 2760
QY 2761 GCATTCGAAAGACTTACCGCGGAAAGCGTTATGTGTGACAGTACAAAGTCAAGAGAT 2820
DB 2761 GCATTCGAAAGACTTACCGCGGAAAGCGTTATGTGTGACAGTACAAAGTCAAGAGAT 2820
QY 2821 CCACTACTGCGAGACTTCTGAGCAGTGAACGTTACTTACACGCGAGAAACACG 2880
DB 2821 CCACTACTGCGAGACTTCTGAGCAGTGAACGTTACTTACACGCGAGAAACACG 2880
QY 2881 ATTGTCTGGAAGACGCTAGCTGTGTATCCGTGATTAAGACACTTACGCTTAAATCC 2940
DB 2881 ATTGTCTGGAAGACGCTAGCTGTGTATCCGTGATTAAGACACTTACGCTTAAATCC 2940
QY 2941 GGGGATTTACAGGCTTCTTGAAGACAGTGGAGCGGCAACGAGCCATTAATGCGC 3000
DB 2941 GGGGATTTACAGGCTTCTTGAAGACAGTGGAGCGGCAACGAGCCATTAATGCGC 3000
QY 3001 GTTCTGTATAGCCGACAGACGCTGATGTGTTCCAGAAATAGGTGAACGTGCGTGGCG 3060
DB 3001 GTTCTGTATAGCCGACAGACGCTGATGTGTTCCAGAAATAGGTGAACGTGCGTGGCG 3060
QY 3061 AAGGCTTTAGACAGTCTTGGGCAACGCTGATTTGCTGAGAGACAGACAGTGGAG 3120
DB 3061 AAGGCTTTAGACAGTCTTGGGCAACGCTGATTTGCTGAGAGACAGACAGTGGAG 3120
QY 3121 ACGTTGCAACCATTCAGACATGACAGCGGTACTCACTGAAATGCGACTGAACCTTCTT 3180
DB 3121 ACGTTGCAACCATTCAGACATGACAGCGGTACTCACTGAAATGCGACTGAACCTTCTT 3180
QY 3181 TGCAACAGGCTCTTGTGAGTGAAGCTGAGCAGTGGGTTATTTCCGCTCCATACGCTGCA 3240
DB 3181 TGCAACAGGCTCTTGTGAGTGAAGCTGAGCAGTGGGTTATTTCCGCTCCATACGCTGCA 3240
QY 3241 CTTAATTAACAGGATTCAGACATGAGATTAATCTCGCAGGGAAGAACATGTATGAGCTTAAT 3300
DB 3241 CTTAATTAACAGGATTCAGACATGAGATTAATCTCGCAGGGAAGAACATGTATGAGCTTAAT 3300
QY 3301 AGAAGGTAGCAAGAGGATTTGTCAAGCGGATATCCGTGATCAACAAAGCGGTTGACACA 3360
DB 3301 AGAAGGTAGCAAGAGGATTTGTCAAGCGGATATCCGTGATCAACAAAGCGGTTGACACA 3360

QY 3361 GGCAGGATGCTGATATTAAGAAATTAATCAATCAAGGACTACTCTCCAAATTAATGTG 3420
DB 3361 GGCAGGATGCTGATATTAAGAAATTAATCAATCAAGGACTACTCTCCAAATTAATGTG 3420
QY 3421 GTTCCATTAATTCGCGGTTGCCCACTGTTGATCGTTGACCAAAAGACAGGGTACA 3480
DB 3421 GTTCCATTAATTCGCGGTTGCCCACTGTTGATCGTTGACCAAAAGACAGGGTACA 3480
QY 3481 ACTGATCAACGCGGATTCCTATCTAAGATGAAGGCAAAATGTGTTGTGTATGCGGAT 3540
DB 3481 ACTGATCAACGCGGATTCCTATCTAAGATGAAGGCAAAATGTGTTGTGTATGCGGAT 3540
QY 3541 CCTATCAGCATTCAGAGGAAGAAAGTAAAGTCAATGAGTCCATGCTCCATTAATCAATC 3600
DB 3541 CCTATCAGCATTCAGAGGAAGAAAGTAAAGTCAATGAGTCCATGCTCCATTAATCAATC 3600
QY 3601 AGGTGTATCTGATTTGGGAATACTAAGCATGTGCGTAATAATGAATTAATCTTTGTC 3660
DB 3601 AGGTGTATCTGATTTGGGAATACTAAGCATGTGCGTAATAATGAATTAATCTTTGTC 3660
QY 3661 AATGTTAGACCCCGCTACAGAAACCATCACTAACCAAGTGGCGAGATCAAGCTTACAC 3720
DB 3661 AATGTTAGACCCCGCTACAGAAACCATCACTAACCAAGTGGCGAGATCAAGCTTACAC 3720
QY 3721 CACAGCATGCTTAAGGTGTAAGGCTGTCCACCACTGAAACATGCGGGAACATGTGTGCT 3780
DB 3721 CACAGCATGCTTAAGGTGTAAGGCTGTCCACCACTGAAACATGCGGGAACATGTGTGCT 3780
QY 3781 ATAGGGTATGCGCTTGTGTATGCGGCAACGAGAAATATCATCTGCGGTGACAGCTCA 3840
DB 3781 ATAGGGTATGCGCTTGTGTATGCGGCAACGAGAAATATCATCTGCGGTGACAGCTCA 3840
QY 3841 TTTAGGTTTACCCGTGTGTGTGACCTTAAGAACTGCGGAAATATCTGAGGTTCTCTTC 3900
DB 3841 TTTAGGTTTACCCGTGTGTGTGACCTTAAGAACTGCGGAAATATCTGAGGTTCTCTTC 3900
QY 3901 GTGTTCTTGGCAAGGACAACGCGCAACCAATGACATGACAGGACAGCTGCGGTGATG 3960
DB 3901 GTGTTCTTGGCAAGGACAACGCGCAACCAATGACATGACAGGACAGCTGCGGTGATG 3960
QY 3961 CTGACAAATCTATCAAGGCTCAACAGGTAACGAGGACAGGAGCTCCAGCGTACAGA 4020
DB 3961 CTGACAAATCTATCAAGGCTCAACAGGTAACGAGGACAGGAGCTCCAGCGTACAGA 4020
QY 4021 GTGATCAAGAGTGAATTAAGCAAGGCGCTGACCAACTATCTTAAATGCTGCTAATGAC 4080
DB 4021 GTGATCAAGAGTGAATTAAGCAAGGCGCTGACCAACTATCTTAAATGCTGCTAATGAC 4080
QY 4081 AAAGGTCAACGAGTTCCGAGTGTGCGGTGCACTGTACGAAATGCGCGGCTGCTTT 4140
DB 4081 AAAGGTCAACGAGTTCCGAGTGTGCGGTGCACTGTGTACGAAATGCGCGGCTGCTTT 4140
QY 4141 GATAGACAGCAATAGCTGTGCGGACCGCTAGACTTGTGAAGCAGAACCGCTCATCATA 4200
DB 4141 GATAGACAGCAATAGCTGTGCGGACCGCTAGACTTGTGTGAAGCAGAACCGCTCATCATA 4200
QY 4201 CATGCTGTAGAGCCCAATTTTCTAAGATGCGGAAACCGAGGCGCACTTAAGCTGCA 4260
DB 4201 CATGCTGTAGAGCCCAATTTTCTAAGATGCGGAAACCGAGGCGCACTTAAGCTGCA 4260
QY 4261 GCTGCTTACATGAGCATAGGCTCATGCTGTAACGCTGAGCGGATTAACAAATATCAAGTA 4320
DB 4261 GCTGCTTACATGAGCATAGGCTCATGCTGTAACGCTGAGCGGATTAACAAATATCAAGTA 4320
QY 4321 CCGCTACTGTCAACCGGCACTTATCTGTGTGCAAAAGATGAGTATCATTCAT 4380
DB 4321 CCGCTACTGTCAACCGGCACTTATCTGTGTGCAAAAGATGAGTATCATTCAT 4380
QY 4381 CACTGTTCACCTGTTTGCACATGACGATGCGGATGTCACTATATTTGTTGGATTA 4440
DB 4381 CACTGTTCACCTGTTTGCACATGACGATGCGGATGTCACTATATTTGTTGGATTA 4440
QY 4441 CAATGGAGACAGGATTAATGAGGCAATTCACCGCAAAAGAACGCTGAAATTCGTGAT 4500

Db 4441 CAATGGAGACACGATATATGAGGCGCTTCAACGCAAGAAAGCTCGAAATTCGTGAT 4500
Qy 4501 GATACAGACGACGATAGACATTTGATGAGGATCCACCCAAACACTCTTTGGCAGGC 4560
Db 4501 GATACAGACGACGATAGACATTTGATGAGGATCCACCCAAACACTCTTTGGCAGGC 4560
Qy 4561 AGACAGAGTTACTCTGCTCAATGAGGCAAGTTGATTCACTCTGAGAGTACAGATTC 4620
Db 4561 AGACAGAGTTACTCTGCTCAATGAGGCAAGTTGATTCACTCTGAGAGTACAGATTC 4620
Qy 4621 CATCAGACCGGCAAGACATTTGCCAAATCCATCAATGTGGCCCAAAATCTGAGGCT 4680
Db 4621 CATCAGACCGGCAAGACATTTGCCAAATCCATCAATGTGGCCCAAAATCTGAGGCT 4680
Qy 4681 AATGACGAGATTTCTGTTGATCATCTGGGGAGAGTATGTCCAGCATCCGCTCCAAATGC 4740
Db 4681 AATGACGAGATTTCTGTTGATCATCTGGGGAGAGTATGTCCAGCATCCGCTCCAAATGC 4740
Qy 4741 CCAATAGAGAGTACAGAGGCGTCTGCTCACTCAACACTTCATGCGCTGTATATAC 4800
Db 4741 CCAATAGAGAGTACAGAGGCGTCTGCTCACTCAACACTTCATGCGCTGTATATAC 4800
Qy 4801 GCTATGACGCGTGAACGCGTATACAGGTTGCGCTCTGGAGAAAGAAACAGTTGCGCGTA 4860
Db 4801 GCTATGACGCGTGAACGCGTATACAGGTTGCGCTCTGGAGAAAGAAACAGTTGCGCGTA 4860
Qy 4861 TGCTCATCATCTCTGTTGCGGAAATACAGATTCACAGCGCTGCGAAAGCTACATGACG 4920
Db 4861 TGCTCATCATCTCTGTTGCGGAAATACAGATTCACAGCGCTGCGAAAGCTACATGACG 4920
Qy 4921 AAACAGCTCTGTTTACAGGCGTGTACCAACGCGCTGTACACCCAGAGAAATACGCGAA 4980
Db 4921 AAACAGCTCTGTTTACAGGCGTGTACCAACGCGCTGTACACCCAGAGAAATACGCGAA 4980
Qy 4981 ATAAATCTAGAAACGCCACC 5000
Db 4981 ATAAATCTAGAAACGCCACC 5000

RESULT 2

US-09-991-258-1

Sequence 1, Application US/09991258
Patent No. US20020141975A1

GENERAL INFORMATION:

APPLICANT: Olmsted, Robert

APPLICANT: Keith, Paula

APPLICANT: Dryga, Sergey

APPLICANT: Caley, Ian

APPLICANT: Maughan, Maureen

APPLICANT: Johnston, Robert

APPLICANT: Davis, Nancy

APPLICANT: Swanstrom, Ronald

TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE

FILE REFERENCE: 0113.0001U3

CURRENT APPLICATION NUMBER: US/09/991,258

CURRENT FILING DATE: 2001-11-16

PRIOR APPLICATION NUMBER: 09/902,537

PRIOR FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: 60/216,995

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 12523

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE: Description of Artificial Sequence; No. US20020141975A1 =

OTHER INFORMATION: synthetic construct

US-09-991-258-1

Query Match 42.3%; Score 2116.4; DB 10; Length 12523;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3218; Conservative 0; Mismatches 1746; Indels 24; Gaps 3;

Qy 10 ACTAATGATCCAAATATGAAAGAAATTTACGTTGACTTAATGCTGACACCCGATGTC 69
Db 30 AATTAATCCAAATATGAAAGAAATTTACGTTGACTTAATGCTGACACCCGATGTC 89
Qy 70 AAGTCGTTACAGGAGGAGCTTTCACAAATTTAGATCGAAGACAGGAGCTGACAT 129
Db 90 AGAGCTTTCAGAGGAGGAGCTTTCACAAATTTAGATCGAAGACAGGAGCTGACAT 149
Qy 130 GACCATGCAATGCAAGAGGCTTTCGATGTCGCAACAAAGCTCAATGAGAGGAGTC 189
Db 150 GACCATGCAATGCAAGAGGCTTTCGATGTCGCAACAAAGCTCAATGAGAGGAGTC 209
Qy 190 GACCGGAGCAAGTTATCTTGAACATTGGAAGTGGCCGCTCAAGATGCAATTCCAAT 249
Db 210 GACCCATCCGACAGATCTTGAACATTGGAAGTGGCCGCTCAAGATGTAATTTCTAG 269
Qy 250 CACCGTATGATGTAATCTGCGCTATGATGTAAGGCTGGAAGCCGGACAGTACAAACG 309
Db 270 CACAGTATGATGTAATCTGCGCTATGATGTAAGGCTGGAAGCCGGACAGTATGTAAG 329
Qy 310 TATGCAAGAAAGCTTAAGAAAGT-----GACATTACCAAGCAAGATAGCCTTAG 363
Db 330 TATGCAAGTAAAGCTTAAGAAAGTAACTGAAGAAATATGATAGAAATTTGCAAGAAA 389
Qy 364 GCGGACAGCTGCTGGAATGATGTCACACAGAGCAGAGACTCATCTCTGTATG 423
Db 390 ATGAAAGAGCTCCCGCGCTGATGAGGACCTGACCTGGAACCTGAGATATGTCCTC 449
Qy 424 CACACAGAGCCACGCTGATGTAATCTTGAAGTGTAGCAATATACCAAGATGTATGCA 483
Db 450 CACAGCAGAGTGTGTCTGCTACGAAAGGCAAGTGTCTTTACAGATGTATGACG 509
Qy 484 GTCCATGCAACGACATCAATCTACACAGGCGCTTAAGAGTTAGCAATTTACTG 543
Db 510 GTTACGAGACGCAAGTCTCTATCAACAGCCAAATGAGAGTTAGTGGCTACTG 569
Qy 544 ATAGCTTTGACACGACCCCTTTTATGTAACAAACATGCGAGTTCTACCTTACTAC 603
Db 570 ATAGCTTTGACACACCCCTTTTATGTAACAACTGCTGAGCATTCATCTATC 629
Qy 604 AACAGAACTGGGCTGACGAGAGATTGGAAGCACTGGAACCTGCGTAATCA 663
Db 630 TCTACCACTGGGCTGACGAAACCGTCTTAACGCTGTGTAACATGAGCTATGACGCT 689
Qy 664 GATCTTCAGAGAGAGGCTTGAAGAACTCTCAATCTTGAAGAGAGGCTCAACT 723
Db 690 GAGCTTATGAGAGGCTGACGTAAGGAGTGTCCATCTTGAAGAGAGTATTTGAAACA 749
Qy 724 ACTAATAGATCATATTTCTGTTGTTCAACATCTACACAGAGATGATCACTGTTA 783
Db 750 TCCAACATCTTCTATCTCTGTTGCTCAGCATCTACACAGAGAGGAGCTTACTG 809
Qy 784 CTAAGCTGAGATCTTCCAACGCTTCACTGAAAGAGAGTCACTCAACAGTAGA 843
Db 810 AGAGCTGAGACCTGCGCTGTATTTCACTTACGTGCAAGCAAAATTTACACATGTCG 869
Qy 844 TGTGAGACATTTGTCAGCTGTAAGGTAAGTATCAAAAGATTAACATGAGCCAGCA 903
Db 870 TGTGAGCTATGTAATTTGTAAGGTAAGTATGTAAGATTAAGATGATCAAGTCCAGGC 929
Qy 904 CTATACGTAAGTGAAGCTTGGCTCAACATGATGCGAGGCTTCTGAGTTGC 963
Db 930 CTGATGGAAGCTTACGCTATGCTGTAAGTACATGACGCGAGGATTTCTGTGCTGC 989
Qy 964 AAATGACATATGCTGCGCGGAGAGGTTCTTCTGTGTATGATGATGCA 1023
Db 990 AAATGACATATGCTGCGCGGAGAGGTTCTTCTGTGTATGATGATGCA 1049
Qy 1024 GCCACATTTGCTGATGATGATGACAGGATTTGGAACATGACGTTAGTGTAGCA 1083

Db 1050 GCTAATTGTGTGACCAATGCTGCACTACGCGAACAAGATGATGCTGCGGACGACGCG 1109
Qy 1084 CAAAAAATTATGGGCTCAACCAAGGATTTGTGCAATGGTAAAGCGCAAGAAAT 1143
Db 1110 CAAAAAATCTGTGTGGCTCAACGAGCTATATGTGTCAAGCTGCGACCGCAAGAAAC 1169
Qy 1144 ACTAACAATGCAACATATCTATTTACAGTGTGTGCGCCAGCGCTTTTCCAGTGGCG 1203
Db 1170 ACCAATACCATGAATAATTAATCTTTTGTGCGTAGTGGCCAGGCAATTTGCTAGTGGCA 1229
Qy 1204 CGTGAACATCGTGGCACTTGGACGACGAAAGAACTAGGGGTGCGGAGCGCACTCTT 1263
Db 1230 AAGGAATTAAGGAAGATCAAGAAATGAAAGCCACTAGGACTAGAGATGACAGTTA 1289
Qy 1264 ACTATGGGCTGCTGGGCTTTCAAGACCCAGAAATACATCCATCTACAGAAAGCT 1323
Db 1290 GTCAATGGGCTGTTGTGGCTTTAGAAAGGCAAGATTAACATTAATTAAGAGCGCCG 1349
Qy 1324 GGTACGCAACAAATTAAAGAAAGTACCTGCGCTTTGATCTCATTTTGTGATTCAGCGCTT 1383
Db 1350 GATACCCCAACCATATCAAAAGTGAACAGCGATTTCCACTCATTTGTGCTGCCAGATA 1409
Qy 1384 ACCAGCCAGGGCTCGATATGGGCTTCCGCGTAGGCTCAAGCTGCTGTTGAACCACT 1443
Db 1410 GGCAGTAAACACATTTGGAATCGGGCTGAGAACAAAGATCAAGAAATGTTAAGAGAGC 1469
Qy 1444 GTCAACCCGCACTGCTAATTAATGCGCCGATGTGAGACATCTGCGGTTCACAGCA 1503
Db 1470 AAGGCGCGTCACTCTCAATTAACCGCCAGAGCAAGAAAGCTAAGTGCACCGCAT 1529
Qy 1504 GAAGCTGAAGAAAGTGTGCTGACGCGGAGAGATCAGAGAAAGCTTCCACCTTGTCTCT 1563
Db 1530 GAGGCTAAGAGAGTGTGCTGAGACCGAGAGATTTGCGCGAGCTTACACCTTTGGCAGCT 1589
Qy 1564 GAATAGAAAAAGAGACCGTAGAGGCAAGAGTGAACCTTATTATGCAAGGCGAGAGCA 1623
Db 1590 GATGTTGAGAGACCCACTCTGGAAACCGATGCGACTTGAATGTTCAAGAGGCTGGGCG 1649
Qy 1624 GGTAGCGTGAAGACCAAGAGAGACATCAAGGTGACAAAGTTACCGAGCGCAAGAG 1683
Db 1650 GGTCAAGTGAAGACCTCTGTGGCTTGAATTAAGGTTACCAAGTACGTCGCGAGAGCAAG 1709
Qy 1684 ATGGGCTTATGCTATCTTCAACCCAGGCGGATTGAATGTAAGAAATCTGCGTGT 1743
Db 1710 ATCGCTTATGCTGTGCTTCTCCGAGGCTGATCTCAAGAGTGAATAATTAATCTTGC 1769
Qy 1744 ATCCACCCATTGGCGGAACAAGTACTGTATGACTCAAAAGTAAAGGCGAGAGATAC 1803
Db 1770 ATCCACCTCTGCTGAACAAGTCAATGATTAACACTCTGCGCAAAAGGCGGTTAT 1829
Qy 1804 AAAGTCGAGCCATACCAAGGTAAAGTCAATGTAACGAAAGGAGCGGCTCTTTCA 1863
Db 1830 GCCGTGAACCATACATGATGTAAGTGTGTCAGAGGAGCATGTCAATACCCCTCCAG 1889
Qy 1864 GACTTCAGGCAATTGATGAGAGCGCTACGATGTTTTCAAGGAGAGGAGTTGTTAAAC 1923
Db 1890 GACTTTCAAGCTCTGAGTGAAGTGCACCATTTGTCTACAGGAACTGAGTTCTTAAC 1949
Qy 1924 AGATACCTGCAACCATGCAATCAACGAGAGGCGTAAACATGACGAAAGTACTAT 1983
Db 1950 AGGTACTGCAACCATATTGCGCACATGAGAGGCGGTGAACACTGATGAAGAAATATAC 2009
Qy 1984 AAAGCTGAAGAGCTCAAGACACAGACTCAAGAAATGCTTTGATTAATGACGACGAAAG 2043
Db 2010 AAAGCTGAAGAGCCAGAGAGACGACGAGCAATCTGTATGACATGACGAGAAACAG 2069
Qy 2044 TGTGTTAAGCGAAGAGACGAGGTCCCTTGTGCTTAACGCGTATGATCTGTAGATCCACA 2103
Db 2070 TGCCTCAAGAAAGAACTAGTACTGAGGCTACAGGAGAGCTGTGATCTCTCC 2129
Qy 2104 TTTCAGAGTTTGTGCTACGAGTCTCAAGACACGACGAGCACTTCACAAGTCCCA 2163

Db 2130 TTCAATGAATTCGCTACAGAGATCTGAGAACACGACGAGCGCTCTTACCAGTACCA 2189
Qy 2164 ACATCGAGGTCTATGAGTGTGACAGTTCAGGTAAATCTGGAATCATCAAAAGCGCTGTG 2223
Db 2190 ACCATAGGAGTATAGGCTGTCCAGATCAGCAAGTCTGGCATATTAAGAGCCAGTC 2249
Qy 2224 ACTAAGAAAGATCTGTGTGATGTCGAAAGAGAAACCTGCGAGAAATCATCAGGAT 2283
Db 2250 ACCAAAAAGATCTATGTTGTGAGGCGCAAGAAAGAAACCTGTGCAAAATTAAGAGGAC 2309
Qy 2284 GTAAAGAGATGACAGTATGATGTGCTGTAGAGCTGTGATTCAGTGTCTTAAT 2343
Db 2310 GTCAAGAAATGAAGAGGCTGAGCTCAATTCAGAACCTGTGAGTCTGAGTCTTGAAT 2369
Qy 2344 GGGGTTAGCAACCCGTTAAACATCTGATCAATGATGAGGCAATTTGCCGATGACAGG 2403
Db 2370 GGAATCAACACCCGTAAGAACCTGTATATGACGAAGCTTTGCTGTGATCAGAGT 2429
Qy 2404 ACGTCTGAGCACTGATTTGCCATGCTCAAAACCTAAGAAAGTGAATTTGGGGGACCCA 2463
Db 2430 ACTCTCAGAGGCTCATAGCATTAATTAAGACCTAAGAGGAGTCTCTGGGGGATCC 2489
Qy 2464 AAACATGCGGCTCTTAAATGATGTGCTGAAAGTACATTTTAACATGACATATGC 2523
Db 2490 AAACAGTGGGTTTTTTTAAATGATGTGCTGAAAGTCAATTTTAACAGAGATTGC 2549
Qy 2524 ACTGAGTATCACTAATTAAGATCTCTAGAGGTGCAACAGACTGTAAACCGCATGCTC 2583
Db 2550 ACACAGTCTTCAACAAAGATCTCTGCGGTGACATTAATCTGATCTTGGTCTCTC 2609
Qy 2584 TCCAGGCTCTTCTACGACAAAGCAAGTGAAGACGTTAAACCATGCTGATTAATATC 2643
Db 2610 TCAACCTGTTTACGACAAATAATGAACGACGAATCCGAAGAGACTAAGTTGTG 2669
Qy 2644 ATGATATCACAGGACCAACAAAGCCGCAAAAGATGCTGATTTCAACCTGTTTCA 2703
Db 2670 ATTGACATTAACCGGACAGTACCAACCTTAAGCAGACGATTCATCTCACTTGTTCAGA 2729
Qy 2704 GGAATGGTGAACAGCTACATTAATGATCAACAAATCAAGAAATATGATCTGCGCTCA 2763
Db 2730 GGGTGGTGAAGCAGTGTCAATGATTAACAAAGGCAACGAAATATGACGAGCTGCC 2789
Qy 2764 TCGCAGGACTTACGCGGAAAGGCGTTATGCTGTAGAGTCAAAAGTCAACGAGATCCA 2823
Db 2790 TCTCAGGAGCTGACCCGTAAGTGTGTATCCGTTGCTGATCAAGTGAATGAATCTCT 2849
Qy 2824 CTCTAATGCACTTGTGACACGATGAACGTTTACTTACACGACGAAAAACGACT 2883
Db 2850 CTGTACGCAACCACTTCAAGACATGTGAACGTTCTACTGACCCGCAAGAGGACGCAATC 2909
Qy 2884 GTCTGAGAGAGCTATGCTGTGATTCCTGTGATTAAGACATCTTACAGCTTAATATCCGGG 2943
Db 2910 GTGTGAAAAACATGACCGGAGCCATGATTAAGAAACATGACTGCCAATCCCTGGG 2969
Qy 2944 GATTTCAGGCTTATATGAGACGACTGAGCGGCGCAACGACGATTAATGAGCAAGCTT 3003
Db 2970 AATTTCAGTGCAGATTAAGAGATGTCACAGAGCTATGATGATTAAGGACATC 3029
Qy 3004 CTGTATTAAGCCGACAGACAGTGTATGTTCCAGATTAAGTGAAGTGTGCTGGCGAAG 3063
Db 3030 TTGAGAGACCGGACCTTACCGAGCGTCTTCCAGATTAAGGAAAGTGTGTTGGCCAAAG 3089
Qy 3064 GCTTTAAGCCAGTCTTGTGACCGCCCAACATTTGTGCTGACAGACAGCACTGGAGAGC 3123
Db 3090 GCTTTAAGCCAGTCTTGTGAAACCGCTGACATGATGATGATGATGATGATGATGATGAT 3149
Qy 3124 TTGACCACTTCAAGATGACAGAGCTTCACTCACTGAATGAGCACTGAACCTTCTTTT 3183
Db 3150 GTGATTAATTTTGAAGAGGACAAAGCTTCACTCAGCAGAGATGATGATGATGATGATGAT 3209
Qy 3184 ACCAGTTCCTTGAAGTGAAGCTTGAACAGTGGTTATTTTCCGCTTCAACCTGCGACTT 3243
Db 3210 GTGAGGTTCTTGAAGTGAAGCTTGAAGCTGAGCTCGGCTATTTTCTGACCACTGTTCCGTTA 3269

; SEQ ID NO 14
; LENGTH: 12379
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-14

Query Match 42.3%; Score 2113.2; DB 10; Length 12379;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3216; Conservative 0; Mismatches 1748; Indels 24; Gaps 3;

10 ACTAATCCATCCAAATATGGAAGAATTCACGTTGACCTTAGATGTGACGCCGCTATGTC 69
30 AATTACCTAACCAAAATGGAAGAAAGTTCACTGACATGAGAGAAAGACGCCCATCTCTC 89
70 AAGTCGTTACAGCGGACGTTTCCACAATTGAGATCGAAGCAAGGCGAGTCACTGACAA 129
90 AGAGCTTTGACAGCGAGCTTCCCGAGTTTGAGTAGAAGCCAGACAGTCACTGATATAT 149
130 GACCATGCCAATGCGAGCGCTTTTCCGATGTCGCAACAAAGCTCATTGAGAGCGAAGTC 189
150 GACCATGCTAATGCTCAGAGCGTTTTCGATCTGGCTTCAAACTGATCGAAACGAGGTG 209
190 GACCGGACCAAGTTATCTTGACATTTGGAGATGGCGCCGCTCAGACATGACATCCAA 249
210 GACCATCCGACACGATCTTGAACATTTGAGATGGCGCCCGCCGAGAAATGATTTCTAG 269
250 CACCGCTATCATTTGATCTGCGCTTATGATTAAGCGGTGAAGACCCGGAACAGACTCAAC 309
270 CACAAGTATCATTTGATCTGCGCTTATGATTAAGCGGTGAAGACCCGGAACAGACTCAAC 329
310 TATGAGAAAGACTTAAGAAAAGT-----GACATTACCGACAAAGAACTAGCCTCTAG 363
330 TATGCACTAAGCTGAAGAAAAGCTGAAGAAATTAAGTAAAGAAATTTGACAAAGAA 389
364 GCGGACAGCTGCTGGAAGTCACTGCAACACGAGCGGAGAGACTCCATCTGCTGATG 423
390 ATGAAGAGCTGCGCCGCTCATGAGCGACCTGACCTGGAACCTGAGACTATGAGCTTC 449
424 CACACAGACGCGACGTTGAGTACTTTGGAAGTGAAGAGTATACCAAGATGTTACGCA 483
450 CACGACGACGAGCTGCTGCTGCTACGAGGCGCAAGTGGCTTTACCAAGATGTTACGCG 509
484 GTCCATGACCGGACATCAATCTACCAAGGCGCTTAAAGAGATTAGCAATTTACTGG 543
510 GTTGACGAGCGGACAACTCTATCAACCAAGCAATTAAGGAGTTAGAGTCCCTACTGG 569
544 ATAGCTTTGACAGACCGCTTTTATGTACAAAACATGGCGAGTTCCCTACCTACTTAC 603
570 ATAGCTTTGACAGACCGCTTTTATGTAAAGAACTGGCTGGGACATATCATCATACAC 629
604 AACAGCACTGGGCTGAGAGAGATATTTGAAGACAGTAACTATGGCTCGCTGTAATCA 663
630 TCTAACCACTGGGCGGAGGAAACGATGTTAACGCTCTGTAACATAGGCTTATGACGCTCT 689
664 GATCTTCAGAGAGAGGCTTGGAAAACCTTCAATCTTTAGAGAAAGAGGCTCCAACT 723
690 GACGTTTGAAGAGCGCTCACTGAGAGGATGTCATTTTGAAGAAAGATTTTGAACCA 749
724 ACTAATAGATCATATTTCTCGGTTGGTTCAACATCTACAGAGATAGATCACTGTTA 783
750 TCCAACTAATGTTCTATCTCTGTTGCTGACCACTTACACGAGAGAGGAGGAGCTTACG 809
784 CGTAGCTGGCATCTTCCAAACGTTGTTCACTTGAAGGAAAGTCTTAATCTTACAGAGTGA 843
810 AGGAGCTGGACCTGCGCTGCTGATTTTCACTTACGTTGGAAGCAAAATTTACCAATGTTGG 869
844 TGTGGACCATTTGTCACTGTGTAAGGTTACGTTCACTCAAAAGATTAACGATGAGCCAGGA 903
870 TGTGAGACTATATGTTAGTTGCGACGAGTACGTCGTTAAAGAAAGATGCTACAGTCCAGGC 929

904 CTATACGGTAAAGTTGAGAACTTGGCTCCACAAATGATCGGAGGCTTTCTTGAATGTC 963
930 CTGTATGGGAAGCCCTTACGGGCTATGCTCTAGATGACCCGGAGGAAATTTCTGTGCTGC 989
964 AAAGTCAACAGATACCGCTCGCGGCGAGAGGCTTTCTTTGCTGTGTGTAACGATGTACCA 1023
990 AAAGTACAGACACTTCAACGCGGAGAGGCTCTTTTCCGCTGTGACGATGTGCA 1049
1024 GCCACACTTGGGATCAGATGACAGGAGATTTGCGCACTGACGTTAGTGTGATGACGCA 1083
1050 GTTACATTTGTGACCAATATGACTGCACTGCTGCAACAGATGTGATGCGGACGACGCG 1109
1084 CAAAACTATTTGTTGGCTCAACCAAGAGATTTGCTCAATGTGTAGGACGCAAGAAAT 1143
1110 CAAAAAGCTGCTGGCTCAACCAAGATTTAGTGTGCAACGCTGCAACCCAGAGAAAC 1169
1144 ACTAACACATGACAGAACTATCTATTAACAGTGTGCGCCAGGCGTTTCCAGTGGGCG 1203
1170 ACCAATACCATGAAATAATTTACTTTTGGCGTAGTGCCGAGGCAATTTGCTAGTGGGCA 1229
1204 CGTGAACATCGTGCGCACTTGAACGACGAGAAAGAACTAGGAGTGGCGGAGCGCACTCT 1263
1230 AAGGAATTAAGAAAGATCAAGAGATGAAGGCCACTAGACATGAGATGACAGTTA 1289
1264 ACTATGGCTGCTGCGGCTTTCAAGACCCAGAAATCAGATCCATCTACAAAGGCT 1323
1290 GTCAAGGAGTGTGTGGCTTTTGAAGGACAAAGATTAATCTATTTTAAGGCGCCG 1349
1324 GGTACGCAACAAATTAAGAAAGTACCTGCGCTTTTGACTCATTTGATTTCAAGCCTT 1383
1350 GATACCCCAACATCATCAAGTGAACAGCATTTCCACTATTTGCTGCGCCAGAGATA 1409
1384 ACCGACCGGCTGATGATGGGCTTCCGCGTAGGCTCAAGCTGCTGTTGAACCACT 1443
1410 GGCAGTAAACATTTGAGATGGGCTGAGAACAAAGTACAGAAATTTTGAAGAGAC 1469
1444 GTCAAAACCGGACCGGCTATTATCAATGCGCGATGAGAGCATCTGCTGCTTACAGCA 1503
1470 AAGGAGCGGTACCTCTCTATTTACCGCGGAGACGTACAAAGATCAATGTCGACGAT 1529
1504 GAAGCTGAAGAGTGGCTGACGCGGAGAGATCAGAGAGCCCTGCAACCTTGTCTCCT 1563
1530 GAGCGTAAAGAGTGGCTGAGAGCGAGAGGATTTGGCGGAGCTTCAACCTTTGGCAGCT 1589
1564 GAATTAAGAAAGAGACCGTGAAGGACAGAACTGATCACTTATTAACAGAGGAGAGCA 1623
1590 GATGTGAGAGCCCACTCTGAAAGCCGATGTGCACTTGATGTTTCAAGAGGCTGGGCC 1649
1624 GTAGCGTGAAGACCAACGAGAGACATCAAGGGTGAACAATTACCCGAGCAAGAGAG 1683
1650 GGTCAAGTGAAGACCACTCTGCTGCTTGAATTAAGTTACCAAGTGAAGAGGAGCAAG 1709
1684 ATTGGCTTTACGCTATATCTTCAACCCGAGCGGATTTGAATAGTGAAGAACTGGCGTGT 1743
1710 ATCGGCTTTACGCTGCTGCTTCTTCCGAGCGTGAATCAAGAGTGAAGAAATTAATCTTGC 1769
1744 ATCCACCATTTGGGGAACAGTACTGTTATGACTCAAAAGTATGGGCGAGGAGATAC 1803
1770 ATCCACCTCTCGCTGAACAAAGTCAATGTTATCAACACTCTGCGCGAAAGGCGGTAT 1829
1804 AAAGTGAAGCATCAAGGATGAGTCAATTTGAACAGAGAGGAGCGGCGCTTGTCA 1863
1830 GCGTGAACATATCAATGTTAGTGTGCTGACAGGAGACATGCAATACCTGCTCAG 1889
1864 GACTTCAGGATTTGATGAGAGCGCTACGATCGTTTCAACGAGAGGAGTTCGTTAAAC 1923
1890 GACTTCAAGCTCTGAGTGAAGAGTGCACACTTGTGTCAACGAACTGAGTGTGTTAAC 1949
1924 AGATTCCTGACCACTATGCAATCAACGAGAGAGCGCTTAACTGACGAGAAAGTACTAT 1983
1950 AGGTACCTGACCACTATTTGCAACATGAGAGAGCGCTGAACACTGATGAAGAAATATTAAC 2009
1984 AAGACTGTAAGACTCAAGACACAGACTCAGAAATACGTTCTGATATTTAGCGCAAGAAAG 2043

Db 2010 AAAAAGTCAAGCCAGGAGACAGAGGCGAATATCTGTACGATGACAGGAAAAG 2069
Qy 2044 TGTGTTAAGCGAAGAAAGACGAGGTCCTTGTGCTTAACCGGTGATCTGTGATTCACCA 2103
Db 2070 TGCCCTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCAAGGCGAGCTGTGTGATCTCC 2129
Qy 2104 TTTCACGATTTGGGTACGAGAGTCTCAAGACGAGCAGCAGGACCTCAAAAGTCCA 2163
Db 2130 TTCCATGATTTCCGTACGAGAGTCTGAGAACAGACGAGCGCTCTTACCAAGTACCA 2189
Qy 2164 ACCATCGAGTCTATGAGTCCAGGTTCAAGTAATCTGAAATCATCAAAAGCGCTGTG 2223
Db 2190 ACCATAGGGGTGTATGGCGTCCAGGATCAGGACAGTGTGCAATCAATTAAGGCGATC 2249
Qy 2224 ACTTAAGAAAGATCTGTGTGTGATGTCGAAAGAAAGAAATCTCCGAGAAATATCAAGGAT 2283
Db 2250 ACCAAAAAGATCTAGTGTGTGAGCGCCAAAGAAAGAAATCTGCAAAATTAAGGAGAC 2309
Qy 2284 GTAGAGGATGAGACGATGATGTTGCTGTAGACTGTGCAATTCAGTCTTCAAT 2343
Db 2310 GTCAAGAAATGAAAGGCTGAGCTCAATCCAGAACTGTGACTGTGCTCTTGAAT 2389
Qy 2344 GGGGTTAAGCAACCCGTTAACTCTGTACATGTAGAGCATTTGCTGCATGCAAGG 2403
Db 2370 GGATGCAAAACCCCGTAGAGACCTGTATATTGACGAAGCTTTTGTGTCAAGAGT 2429
Qy 2404 ACGTGTGCGCATGTATGCTCATGTCAAACCTTAAGAAAGTGTATGTGTGGGGACCA 2463
Db 2430 ACTCTCAGAGCGCTCATATGCTATTAAGAACTTAAGAAAGCGATGCTGTGGGGATCCC 2489
Qy 2464 AAACAATGCGGCTCTTTAATCATGATGTGCTGAAGTACATTTTAACATGACATATG 2523
Db 2490 AAAAGTGGGCTTTTATTAATGATGTGCTGAAGTGTGATTTTAACAGAGATTTGC 2549
Qy 2524 ACTGAAGTGTACATAAAAGCATCTAGAGGTGACACAGACTGTAAACCGCATGTC 2583
Db 2550 ACAAGTCTTCCAAAGCATCTCTGCGTTGCACTAATCTGTGACTTCCGCTGTC 2609
Qy 2584 TCCACGCTCTTCTACGACAGAGCAATGAAGCGTTAAACCATGTGCTGATTAATCATC 2643
Db 2610 TCAACCTTGTTTACGACAAAGAAATGAGACGAGATCCGAAGAGACATTAATGTG 2669
Qy 2644 ATAGATACCAAGGAGACACAAAGCCGACAAAGTATGATTTCACTGTTTGA 2703
Db 2670 ATTACACTACCGGACGATACCAACCTTAGAGAGAGCATCTCATCTTCTCATGTTTCA 2729
Qy 2704 GGATGGGTAAACAGCTACAGATTGACTCAAAATATCAGAAATCATGCTGCGCTGCA 2763
Db 2730 GGGTGGGTGAAGCAAGTTCGAAATGATTAAGAAAGCAAGAAATTAAGACGCGACGTC 2789
Qy 2764 TCGCAAGGACTTACGCGGAAAGCGTTTATGTCTGACGTTCAAAAGTCAAGAGATCA 2823
Db 2790 TCTCAAGGCTGACCCGTAAGGTGTGTATGCGTTCGTAAGAGTGAAGAAATCTCT 2849
Qy 2824 CTCTACTCGGAGACTTGTGACAGAGTGAAGGTGTTACTTACACGACAGAAACGCAAT 2883
Db 2850 CTGTACGACCCACCTCAAGACATGTGAAGTCTCTACGACCGGACGAGAGACGCAATC 2909
Qy 2884 GTCTGAAGAGCTAGTGTGATCCCTGATTAAGACATTAAGCTTAATATCCGCG 2943
Db 2910 GTGTGAGAAACATAGCCGCGGACCAATGATTAAGAAACATGATCGCAATACCTGTGG 2969
Qy 2944 GATTTCACGCTTATTTGAGACGATGCGAGCGGCAACGAGCCATTTATGACAGCGCTT 3003
Db 2970 AATTTCCTGCGCAGATAGAGAGTGGCAACAGAGCATGATGCAATGAGGACATC 3029
Qy 3004 CTTCATTAAGCCGAGACAGCTGATGTGTTCAGAAATTAAGTGAAGTGTGTGGCGAG 3063
Db 3030 TTGAGAGACCGGACCTTACGAGCTTTCAGAAATTAAGGCAAAAGTGTGTGGCGCAAG 3089
Qy 3064 GCTTTAGAGCCAGTCTTGTGGCAGCGCAATTTGTGACGAGACAGCAAGTGGAGAGC 3123

Db 3090 GCTTTAGCCGGGTGCTGAAGACCGTGTGATAGACATGACCACTGAACATGAGCACT 3149
Qy 3124 TTGACCCATTTCAAGCATGACAGAGCTACTCACTGAAATGGACATGAATCTTTGC 3183
Db 3150 GTGATTTATTTGAAAGCAAGCAAGCTCACTACAGCAAGATAGTATTGAACCAATATGC 3209
Qy 3184 ACCAGTTCCTTTGGAGTAGACCTGAGCAGTGGTTATTTCCGCTCTACCGTGCATT 3243
Db 3210 GTGAGTTCCTTTGGACTGATCTGAGCTCGGCTATTTTCTGACCACTGTTCCGTTA 3269
Qy 3244 ACTTACAGGATACGACTGCGGATTAATCCCGCAGGAAAGACATGTTATGCGCTTAATGA 3303
Db 3270 TCCATTGAGATATCACTGGGATTAATCCCGCTGTCAATGTACGGGCTGAATAAA 3329
Qy 3304 GAGGTACCAAGAGTGTGACGCGGATATCCGTCATCAAAAGCGTTGACAGAGC 3363
Db 3330 GAAGTGTGCGTACGCTCTCTCGAGGTACCCACACTGCTGGGAGATGGCCACTGGA 3389
Qy 3364 AGGTAGCTGATTAAGAAATTAATCCATCAAGACTACTCTCCAAATTAATGTGTT 3423
Db 3390 AGAGTCTATGACATGAACATGATGACCTGGCAATTAATATCCGCGCATTAACCTAGTA 3449
Qy 3424 CCATTAATGCGCGTTGCCCACTGTTATGTGACCAAGAACAGGATCAACT 3483
Db 3450 CCGTAAACAGAAAGACTCCTCATGTATTAGTCTCCACCATATAGAACACCCAGAGT 3509
Qy 3484 GATCACAGCGATTCCTATCTAAGTAGAAGGCAAACTGTGTTGTGATCGGATCCT 3543
Db 3510 GACTTTCTTCAATTCGACAGAAATTAAGGCAAGAACTGTGCTGGGTGGGAAAG 3569
Qy 3544 ATCAGACTTCAAGGAAAGAAAGTGAAGTCCATGAGTCCATTAATACATCAAG 3603
Db 3570 TTGTCCCTCCAGGCAAAATGTTGACTGTTGTGACACCGGCTGAGCTACTTCA 3629
Qy 3604 TGTGATCTGATTTGGGAATACCTAGCATGTGCGTAATATGACATTAATCTTTGCAAT 3663
Db 3630 GCTCGCTGATTTAGGCAATCCAGGATGATGTGCCAAATATGACATTAATTTGTTAAT 3689
Qy 3664 GTTAGAGCCCGGTACAGAAACATCACTACCAAGGCGAGGATCAGGATACACAC 3723
Db 3690 GTGAGAGCCCATTAATTAATCACTATCAAGAGTGAAGCCATGCAATTAATCCTT 3749
Qy 3724 AGCATGCTAAGTGTAAAGCTGTCCACCACTGAACTGCGGAAATGATGTGGCTATA 3783
Db 3750 AGCATGTGACAAAGAAAGTGTGTGCTGATCTGAATCCGCGGAACTGTGTGACATA 3809
Qy 3784 GGGTATGGGCTTGTGATGCGGCAACGAGAAATATCATCTGCGGTGACGCTCATTT 3843
Db 3810 GGTATGTATACGCTGACAGGGCCAGGAAAGATCATTTGTGTATAGCGCGCACTTC 3869
Qy 3844 AGTTTACCCGTTGTCTGACGCTTAAGACATGCGGAAATCTGAGGTTCTCTTGCTG 3903
Db 3870 AAGTTTCCCGGTATGCAACCGAAATCTCACTTAAGAGAGCGAAGTTCGTTGTA 3929
Qy 3904 TTCTTCCGCAAGGACCAAGGCAACACACATGACAGGACAGACTCGGTGATGCTT 3963
Db 3930 TTCAITGGATGACATGCCAAGGCCGTGACGACATCTTACAAAGCTTTATCAACCTTG 3989
Qy 4024 ATCAGAGTGAACATTAAGACAGCGCTGACCAAGCTATCTTAAATGTGTGTAATAGAAA 4083
Db 4050 GTGGAAGGAGATTTGCAAGGCGACAGGAGGTGATTAATGATGTGTAACAGAAA 4109
Qy 4084 GGTCAACAGATTCGGAAGTGTGCGGTGACATGACGAATAATGCGCGCTGCTTTGAT 4143
Db 4110 GGAACACTGTGGCGAGGAGGTGTGCGAGCGCTGTATTAAGAAATTCGCGAAGAGCTTGAT 4169
Qy 4144 AGACAGCAATAGCTGTGTGGAGCGGCTAGACTTGTGAAGACGAAAC--GCTCATATA 4200
Db 4170 TTACAGCGATTCGAATGAGAAAGCGGACGACTGTCAAAAGTGCAGCTAAACATATCAT 4229

4201 CATGCTGAGACCAATTTTCTAAGATGCCGAACCGAGGCGACCTTAAGCTCCCA 4260
4230 CATGCCGAGAGCAAACTTCAACAAAGTTTCGAGGTTGAAGAGTGAACAAAGTTGCA 4289
4261 GCTGCTCAATGAGCATAGCTTCATCGTCAACGCTGAGCGGATTAACAAATATCAGTA 4320
4290 GAGGCTTAATGAGTCCATGCTAAGATGTCAACGATTAACAAATGATGAGTCCGATT 4349
4321 CCGCTACTGTCACACCGGATCATCTGCTGGCAAGATTCGATGATCATCATTTGAT 4380
4350 CCACTGTTGTCACCGGATCTTTTCGGGAACAAAGATCGACTAACCAATCATTTGAC 4409
4381 CACCTGTTCACTGCTTTGCAACCTACGATGCCGATGTCACCATATATTTGCTGATAA 4440
4410 CATTGCTGACAGCTTTTGAACCACTGATGAGTATGACCATATATCTGACGGACAG 4469
4441 CAATGGAGACCAAGATATGAGGCAATTCACCGCAAAAGAAAGCGTGAATTTGGA- 4499
4470 AAATGGAAATGACTCTCAAGAGACAGTGGCTAGAGAGAACGATGAGAGATATGC 4529
4500 -----TGATGACAGCCAGTACATTCATTCCTGCTAGGCTCCACCCAAAC 4545
4530 ATATCCGACGACTCTTCAGTGACAGACCTGATGACAGCTGATAGGCTGATCCGAA 4589
4546 AGCTTTTGGAGGACAGACAGGTTACTCCGTCATATGAGGCAAGTTGATTCATACCTG 4605
4590 AGTTCTTTGGCTGAGAGAGGCTACAGACACCAAGTGGCAAACTTTCTCATATTTG 4649
4606 GAAGGTACAGATTCATCAACCGCAAGACATTCGCGAAATCATGCAATGTGGCC 4665
4650 GAAGGACCAAGTTTCAACGAGCGGCAAGATATAGCAAAATTAATGCCATGTGGCC 4709
4666 AACAAATCGAGGCTAATGAGAGATTTGTTTCTATCTTCGGGGAAGATATGTCAGC 4725
4710 GTTGCAAGGAGGCAATGAGAGATGATGATATCTTCGGAAGAAAGATGAGCAGT 4769
4726 ATCCGCTCAATGCGCCAGTAGAGAGGACAGAGGCTGCTCCACCTCAACATTTCCA 4785
4770 ATTAGTGTGAATGCGCCGTCGAGAGAGTGGAGGCTCTCAACCACTGAGCAGTGGCT 4829
4786 TGCTGTGTAATTAATGATGACGCTGAGCGGTATACAGGTTGCGCTTGGCAAGAA 4845
4830 TGCTGTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4889
4846 GAACAGTTCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4905
4890 GAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4949
4906 AAGCTACAGTCAAGCAACAGTCTGTTTTCAGGCTGATGACCAAGGCTGATACAGCC 4965
4950 AAGATCCAAATGCTCCCAAGCTTATTTTCTACCGAAAGTCCGCTATATTTATCA 5009
4966 AGGAAGTA 4973
5010 AGGAAGTA 5017

RESULT 4
US-09-991-258-17
; Sequence 17, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmedo, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swastrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES

FILE REFERENCE: 01113.000103
; CURRENT APPLICATION NUMBER: US/09/991,258
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 17
; LENGTH: 13584
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-17
Query Match 42.3%; Score 2113.2; DB 10; Length 13584;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3216; Conservative 0; Mismatches 1748; Indels 24; Gaps 3;
10 ACTAATGATCAATATGGAAGAATTACGTTGATGATGCTGACAGCCGCTATGTC 69
30 AATTACCTACCCAAATGGAAGATTCACTTGACATCGAGAAACAGCCCATTCCTC 89
70 AAGTGTACAGCGGAGCTTTCACAAATTGATGATCGAAGAGGAGGAGTCACTGACAT 129
90 AGAGCTTGACAGGAGCTTCCGAGTTTGAAGTGAAGGACCAAGGATGATATAT 149
130 GACCATGCAATGCGAGAGCTTTTTCGATGTCACCAAGCTCATTTAGAGCGAAGTCT 189
150 GACCATGATATGCAAGAGCTTTTTCGATGTCGATGTCGATGTCGATGTCGATGTC 209
190 GACCGGAGCAAGTATTTGATGATGAGAGTGGCGGCTGACATGATGATGATGAT 249
210 GACCATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 269
250 CACCGTATCATTTATGCTGCTATGATGATGATGATGATGATGATGATGATGATGAT 309
270 CACAAATCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 329
310 TATGCAAGAACTTAAAGAAAGT-----GACATTAACCAAGAAATGATGATGATGAT 363
330 TATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 389
364 GCGGAGCTGCTGGAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 423
390 ATGAAGGAGCTGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 449
424 CACACAGAGCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
450 CACAGAGAGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 509
484 GTCCATGACAGGATCATCATCAACAGGCGCTTAAAGAGTTAGAGCAATTTACTG 543
510 GTTACGAGCCGACAGATCTATCAACCAACCAATTAAGAGTTAGAGTGGCTACTG 569
604 AATGAGCTTTGACAGAGCCCTTTTATGATCAAAAAACATGAGGTTCTTACTTAC 603
570 ATAGGCTTTGACACACCCCTTTTATGATTAAGAACTTGGCTGAGCATATTCAT 629
604 AATGAGCTTTGAGAGGCTTGAAGAGTATGAGAACTTATGAGTGGCTGATGATGAT 663
630 TCTACCAATGAGGCGGAGAAACGTTGTAACGCTGTAACATAGGCTTATGAGCTCT 689
664 GATCTTGAAGAGAGAGGCTTGAAGAACTTCAATCTTTAGAGAGAGGCTTCAACT 723
690 GACGTTATGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 749
724 ACTAATAAGATCATATTTCTGTTGTTCAACAAATGATGATGATGATGATGATGAT 783
750 TCAACCAATGTTATTTCTGTTGCTGACCATTAACGAGAAAGGAGGACTTACTG 809

QY 784 GGTAGCTGGCATCTTCCAAAGGTGTCCTGAAGAAAGTCTAATTACAGGTAGA 843
DB 810 AGGAGCTGGACCTGGCTCTGTATTACTTAAGTGGCAAGAAAATTACAGTTCGG 869
QY 844 TGTGGGACCATTTGTAGCTGTGAAGGGTACGTCACTCAAAAAGATTAAGCATGCCAGGA 903
DB 870 TGTAGACTATAGTTAGTTGGGACGGGTACGTTCGTAAAGAAATAGCATCAGTCCAGGC 929
QY 904 CTATACGGTAAAGTTGAGAACTTGGCGTCCAAATGCATGCCGAGGGTTTCTTGAATTGC 963
DB 930 CTGTATGGGAAGCCCTTCAAGCTATGCTGTACGATGCAACCCGAGGGATTCCTTGCTGC 989
QY 964 AAAGTCACAGATAGCTGCGCGGAGAGGGTTTCTTTGCTGTGTAGATGATGCA 1023
DB 990 AAAGTCACAGATAGCTGCGCGGAGAGGGTTTCTTTGCTGTGTAGATGATGCA 1049
QY 1024 GCCACACTTTCGATCGATGATGACAGGAAATCTGGCAACTGATGATGATGATGCA 1083
DB 1050 GCTACATTGTGTGACCAATGACTGGCATCTGGCAACAGATGTCAGTGGGAGAGCGCG 1109
QY 1084 CAAAACCTATTGTTGGGCTCAACCAAGATTTGTCGCAATGTTAGAGCCCAAGAAAT 1143
DB 1110 CAAAACCTGCTGTGGGCTCAACAGGCTATGCTGCAACGGTCCGACCCAGAGAAAC 1169
QY 1144 ACTAACA CAATGCAAGAACTATCTATTACCAATGCTGCGCCGAGGGTTTCCAGGTGGCG 1203
DB 1170 ACCTATACATGAAATAATTAATCTTTTGCCGTAGTGGCCAGCAATTTGCTAGGTGGCA 1229
QY 1204 CGTGAACATCGTGGCGACTTGGAGCAGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTT 1263
DB 1230 AAGGAATATTAAGAAAGATCAAGAAATGAAAGGCCACTAGAACATGAGATAGACGTGA 1289
QY 1264 ACTATGGGCTGCTGCTGGGCTTTCAAGACCCAGAAATCATCATCTTACAAGAGCTT 1323
DB 1290 GTCATGGGTGTGTGTGGCTTTTGAAGGCAAGATATCATCTATTATTAAGCGCCCG 1349
QY 1324 GGTACGCAAA CAATTGAAGAAAGTACCTGCGTCTTTCATTTGATGATCCAGCGCTT 1383
DB 1350 GATACCCAAACATATCATTAAGTGAACAGCAATTTCCATCTATGCTGCTCCAGGATA 1409
QY 1384 ACCAGCACA CGGCTCGATATGAGGCTTCCGCGTAGAGCTCAAGCTGCTGTGAACCACT 1443
DB 1410 GGCAAGTAACAATTGGAGATCGGGCTGAGAACAAAGAAATGTTAGAGAGAGAC 1469
QY 1444 GTCAAAACCCGACCGGCTATTACAATGSCCGATGTCAGATCTGCGTGTACAGCA 1503
DB 1470 AAGGAGCCGTCACTCTCATTAACGCGCGAGGACGTACAAAGAACTAAGTGGCAGCGAT 1529
QY 1504 GAAGCTGAAGAGTGGCTGACGCGGAAGATCAAGAAAGCCCTGCGACCTTCTCCCT 1563
DB 1530 GAGGCTGAAGAGTGGCTGAGAGCCGAGAGGTTGCGCGAGCTTACACCTTTTGGCAGCT 1589
QY 1564 GAAATAGAAAAAGAACCGTAGAGGACAGAACTGATCATTTATGCAAGGACAGAGCA 1623
DB 1590 GATGTTGAGAGCCCACTCTGGAAGCGATGTCGATGATGTTACAAGAGGCTGGGCC 1649
QY 1624 GGTAGCGTGGAGACACCAAGAGACATCAAGGTTGACAACTTACCCAGGCGAAGAG 1683
DB 1650 GGCTCAGTGGAGACACCTCGTGGCTTGAATAAGGTTACCACTACGATGGGAGAGACAG 1709
QY 1684 ATTGGGCTTACGCTATCTTCAACCCGAGGCGGATTTAGATAGGAAAACTGGCGGT 1743
DB 1710 ATCGGCTCTTAAGCTGTGCTTTCTCCGAGGCTGTACTCAAGAGTAAAAATTAATCTTGC 1769
QY 1744 ATCCACCATTTGGCGGAACAAGTACTGTTATGACTCAAAAGGTAGGCGAGAGATAC 1803
DB 1770 ATCCACCTCTCTGGTGAACAAGTCAATGATTAACAACCTCTGGCCGAAAAAGGCGTTAT 1829
QY 1804 AAAGTCGACATACCAAGGTAAAGTCAATTTGACCAAGGAGAGCGCGGTCTTGTTCAA 1863
DB 1830 GCCGTGAAACCATACATGATGAAGTGTGTCAGAGGAGACATGCAATACCGTCCAG 1889

QY 1864 GACTTCCAGGATTTAGTGAAGGCTACAGATCGTTTCAACGAGAGAGATTGTAAC 1923
DB 1890 GACTTTCAAGCTCTGAGTGAAGTGGCCACATTTGTATCAACGAACTGATTTCTTAAC 1949
QY 1924 AGATACCTGCACCAATCGCAATCAACGAGAGCGCTTAAACA CTGACGAAGTACTAT 1983
DB 1950 AGGTACCTGCACCAATTTGTCACACATGAGAGCGCTGAACA CTGATGAAGATATTAC 2009
QY 1984 AAGACTTAAGACTCAGACACAGACTCAGAAATACGTTCTTGATATTTAGACGCAAG 2043
DB 2010 AAAACTGTCAAGCCGACGAGCAAGAGGGAATACCTGTATGACATGACAGAAACAG 2069
QY 2044 TGTGTAAAGGAGAAAGACGAGGTCCCTTGTGCTTAAACCGGTATGATGATGACCA 2103
DB 2070 TGCCTCAAGAAAGAACTATGACCTGAGGCTTACAGGCGAGCTGTGTGATCTTCC 2129
QY 2104 TTTCAAGATTTGCTGACGAGTCTTCAGACACGACGACGACCTCAAAATGCCA 2163
DB 2130 TTTCAATGATTTGCTGCTGAGAGCTGAGAAACGACGACCGCTCTTACCAAGTACCA 2189
QY 2164 ACCATCGGATCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGGCTGTG 2223
DB 2190 ACCATAGGGGTATGAGCGGTGCGAGATGAGCAAGCTGGCATCTTAAGAGCGAGTC 2249
QY 2224 ACTAAGAAAGATCTGTTGTGAGTGCAGAAAGAAATGCGCAGAAATCATCAGGGAT 2283
DB 2250 ACCAAABAAATCTATGTTGTGAGCGCCAGAAAGAAATCTGTGCAAAATTTAAGGAC 2309
QY 2284 GTAAAGAGATGAGACGTATGATGTTGCTGTAGAGCTGTGATTCAGTGTCTTAAAT 2343
DB 2310 GTCAAGAAATGAAAGGGCTGAGAGTCAATGACCAAGCTGTGAGCTGAGTGTGTAAT 2369
QY 2344 GGGGTTAAGACCCCGTTAACCTCTGTACATTTGATGAGGATTTGCTGCCCATGCGGG 2403
DB 2370 GGATGCAAAACCCCGTAGAGCCGTGTATTTGACCAACCTTTTGTGTATCAGAGT 2429
QY 2404 ACGCTGTGCACTGATTTGCAATGTCGAACCTAAGAAATGATTTGTCGGGAGCCCA 2463
DB 2430 ACTTCAAGAGCGCTCATAGCTATTAAGACCTTAAAGGACAGTCTCTCGGGGATCC 2489
QY 2464 AAACATGCGGCTTCTTTAACTATGATGCTGTAAGAACTTAACTTAACTATGACATATGC 2523
DB 2490 AAACAGTGGCTTTTAACTATGATGCTGTAAGAACTTAACTTAACTATGACATATGC 2549
QY 2524 ACTGAATGATCAATAAAGCACTCTAGAGAGTGCACAAGACTGATACCCGCAATGTC 2583
DB 2550 ACACAACTCTTCAACAAAGCACTCTGCGCTGTCCTAAATCTGTGACTTGGCTGTC 2609
QY 2584 TCCAGCTCTTCTACGACCAAGCAATGAAGCGTTAACCATGTGCTGATTAATATCATC 2643
DB 2610 TCAACTTGTTTTACGACCAAAATGAGAACAGCAATCCGAAGAGACTTAAGATTGTG 2669
QY 2644 ATAGATACACAGGAGACCAAAAGCGCACAAAGATGATGATTTCACTGTTTACA 2703
DB 2670 ATTGACACTACCGGCTATCAAACTTAACAGGACGATCTCATCTTCTGTTTACA 2729
QY 2704 GGAATGGTGAACAGCTACGATTTGATCAAAATATCAGAAATCAAGATCGGGCTGCA 2763
DB 2730 GGGTGGTGAAGCACTTGAATTAAGATTAAGGCAAGCAAAATTAAGAGCGAGCTGCC 2789
QY 2764 TCCGAAGGACTTACGCGGAAGGCTTATGCTGTCAAGTACAAAGTCAAGCAATCA 2823
DB 2790 TCTCAAGGCTGACCCGTAAAGGTGTATGCTGTGCTGTGATCAAGGATTAAGAAATCT 2849
QY 2824 CTCTACTCGAGACTTCTGACAGCTGAACGTGTACTTACAGGACAGAAAAAGCATT 2883
DB 2850 CTGTACGACACCACTCAGAAACATGTGAACGTCTTACTGACCGGACGAGAGCCGATC 2909
QY 2884 GTCTGAAGACGCTAGCTGTGATCTCTGATTAAGCACTTACAGTAAATATCCCGG 2943
DB 2910 GTGTGAAAAACATGAGCGGCGACCAATGATTAACACTGTGACCAAGTACCCCTGG 2969
QY 2944 GATTTCAGGCTTCAATTGAGAGACTGACGCGGACAGCAAGCCATTAATGACGCGTT 3003

Db 2970 AATTTCTGTCACATAGAGAGTGGCAAGCAGACATGATCCATCATGAGGCACATC 3029
 Oy 3004 CTGTAAGCCGCGACAGACAGCTGATGTGTTCCAGATTAAGTGAAGTGTGCTGGCGAAG 3063
 Db 3030 TTGGAGAGACCGGACCTTACCGACGCTTCCAGAAATAGGCAAAAGTGTGTTGGCCAG 3089
 Oy 3064 GCTTTAGAGCCAGTCTGGCCACGGCCAACTTGTGCTGACGAGACAGAGTGGAGAG 3123
 Db 3090 GCTTTAGTGGCGGTCTGAGAACCGCTGGCATGACATGACCATGAACTAGAACACT 3149
 Oy 3124 TTGCAACCATTCACATGACGAGAGCTACTGAAATGCGACTGAACTTTCTTTTGC 3183
 Db 3150 GTGATTAATTTTGAACCGAACCAACTCACTCAGAGAGATGATTTGTAACCACTATGC 3209
 Oy 3184 ACCAGGTTCTTTGAGTGAACCTGGAACAGTGGTTATTTTCGGCTCTACCGTGCACCT 3243
 Db 3210 GTGAGGTTCTTTGACTGATTTGACTCCGGTCTTATTTTTCGACCCACTGTTCCGTTA 3269
 Oy 3244 ACTTACAGGGATCAGCAGCTGGGATTACTGCGCAGGAAAGAACATGATATGGCTTAATGA 3303
 Db 3270 TCCATTAAGAAATATCACTGGGATTACTCCCGTGGCTTAATGATGAGGCTGAATAA 3329
 Oy 3304 GAGGTAGCAAGAGATTTGTCAGCGGATATCCGTGCATCAAAAACGGTTGACAGGC 3363
 Db 3330 GAAAGTGTCCGTACACTCTCTCGCAGGATCCACAACTGCTCGGGCAGTTGCCACTGGA 3389
 Oy 3364 AGGGATGCTGATATTAAGAAATTAATCAATCAAGAGCTACTCTCAACAATTAATGTGTT 3423
 Db 3390 AGAGTCTATGACATTAACCTGTGTAACCTGCCAATTTGATCCCGCATTAACCTAGTA 3449
 Oy 3424 CCATTAATTCGCGGTTGCCCATCTGTTGATCTGTGACCAACAAAGCAGGGTCAACT 3483
 Db 3450 CCGTAAACAGAAAGCTGCTCACTGCTTATGCTTCCACCAATTAAGAACCCACAGAGT 3509
 Oy 3484 GATCAGACGGATTTCTTATTAAGATGAAGGCCAAATCTGTGTGATGCGCATCT 3543
 Db 3510 GACTTTCTTATCTGACGAAATTTGAAGGCAAACTGCTGTGTCGGGAGAAAG 3569
 Oy 3544 ATCAGCATTCACAGGAGAAAGTAGATGTCATGAGTGCATTTGCCCATTAATCACTACG 3603
 Db 3570 TTGCTGCTCCAGGAAATGTTGATGTTGTGACACCGGCTGAGGCTTACCTTCAGA 3629
 Oy 3604 TGTGATCTGATTTGGGAATACCTAGCCATGCGTAAATATGACATTAATCTTGTCAAT 3663
 Db 3630 GCTGGCTGATTTAGGCATCCAGTGTATGTCCTCAATATGACATTAATTTGTAAAT 3689
 Oy 3664 GTTAGAGCCCGTACAGAAACCATCACTAACAAAGTGCAGAGATCAGCTATCCAGAC 3723
 Db 3690 GTGAGGACCCCATATAAATATACATCACTACAGAGTGAAGACCATGCTTAAGCTT 3749
 Oy 3724 AGCATTCATACGTTAAGGCTGTCCACCACTGAAACATGCGGGAACATGTGTGCTATA 3783
 Db 3750 AGCATTTGACCAAGAAAGCTTGTGATCTGAATCCCGGGAACCTGTGTGACATA 3809
 Oy 3784 GGGTATGGGCTTGTGATCGCGCAACGAAATATCATCACTGCGGTGACAGCTCATTT 3843
 Db 3810 GGTATGTTAGCTGACAGGCGCAGGAACATCTTGTGTATAGCGCGAGATTC 3869
 Oy 3844 AGGTTTACCCGTGTCTGACGCTTAAGAACATGCGGAAATATGAGGTTCTCTTCTG 3903
 Db 3870 AAGTTTCCCGGGTATGCAAAACCGAATCTCTCACTTGAAGAGACGAAAGTTCTGTTGTA 3929
 Oy 3904 TTCTTGGCAGAGCAAGGCAACACACATGACAGACAGCTGGGTATGCTT 3963
 Db 3930 TTCTATGGGTAGATCGAAGGCCGCTGACCAATCTTACAACTTTCATCAACCTTG 3989
 Oy 3964 GACAACATCTATCAAGGCTCAACAGTACAGGAGGAGAGCTCCAGGCTACAGAGT 4023
 Db 3990 ACGAATCTTATACAGGTTCCAGACTCCAGGAAGCGGATGTGCACCTCATATATGTG 4049
 Oy 4024 ATCAGAGTGAATTAAGCAAGAGCGCTGACCAAGCTATCTTATGCTTAATGACAA 4083

Db 4050 GTGCGAGGGATATTGCCAGCGCAACCGAAGAGTGAATTAATATGCTGTAACAGCAA 4109
 Oy 4084 GATCAACAGGTTCCGAGTGTGCGGTGCTGATCTGTAACGAAATGCGCGGCTGTTGAT 4143
 Db 4110 GACAACTTGCGAGAGGATGTGCGAGCTGTATTAAGAAATTTCCGGAAGCTTCAT 4169
 Oy 4144 AGACAGCAATAGCTGTGCGAGCGGTAGACTTGTGAAGACGAAC---GCTCATATA 4200
 Db 4170 TTACAGCCGATCGAAGTGAAGAAAGCGGATCGGTCAAAAGGTGAGCTAAATATCA 4229
 Oy 4201 CATGCTGAGACCCCAATTTTCTTAAGATGCCGGAACCGAGGCGGACCTTAAGTCCGA 4260
 Db 4220 CATGCCGTAGAGACCAAACTTCAACAAAGTTTGGAGTTGAAGGTGACAAACAGTTGCA 4289
 Oy 4261 GGTGCTCATGAGATGATGCTCCATGCTCAACGCTGACCGGATTAACAAATATCAGTA 4320
 Db 4290 GAGGCTTATGATCTCATGCTTAAGTTTTCACATTAACAAATTAACATCAGTACGAT 4349
 Oy 4321 CCGCTATGTCACCCGATCTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4380
 Db 4350 CCACTGTTGTCCACCGGATTTTTCGGGAAACAAAGATGACATTAACCAATCATTTGAA 4409
 Oy 4381 CACTGTTCACTGCTTTCGACACTAGCGATGCCGATGTCACCATTAATTTGTTGATTA 4440
 Db 4410 CATTTGCTGACAGCTTTAGACACACATGATCAGATGAGCCATATATCTGACAGGACA 4469
 Oy 4441 CAATGGAGACAGGATTAATGAGGCAATTCACCGCAAAAGAGGCTGAAATTTCTGGA- 4499
 Db 4470 AATGGGAAATGATCTTCAAGAAAGAGTGTGAGAGAGAGAGAGATATATGC 4529
 Oy 4500 -----TGATGACAGCCAGTACATTTGCTGAGGGTCCACCCAAAC 4545
 Db 4530 AATCCGACGACTCTTCAAGTACAGAACTGATGAGAGCTGTGAGGTTGATCTCGAAG 4589
 Oy 4566 AGCTTTTGGAGGACGACACAGTTTCTCCGTCAATGAGGCAAGTTGATTAATCACTG 4605
 Db 4590 AGTTCTTGGCTGAGAGAGAGGCTACAGCAAGCATGTCAAAACTTCTCATATTTG 4649
 Oy 4606 GAAAGTACAGATTCATCAGACCGCCAAAGCAATTCGCAAAATCCATGCAATGAGGCC 4665
 Db 4650 GAAAGGACCAAGTTTCAACAGCGCGCCAAAGATTAAGAGAAATTAAGCAATGAGGCC 4709
 Oy 4666 AACAAATCTGAGGCTTAATGACAGATTTGCTTGTATATCTGAGGAGAGATGTCAGC 4725
 Db 4710 GTTGCAACGAGGCGCAATGACAGATATGATGTAATCTCGAGAAAGATGACAGCT 4769
 Oy 4726 ATCCGCTCCAAATCCAGTACAGAGAGTCAAGGCGTGTCTCACTCAGCACTTCCA 4785
 Db 4770 ATTAGTTCGAAATGCCCGTGAAGAGTGGAAAGCTTCTCAGCACTGACGCTGCT 4829
 Oy 4786 TGCCGTGCTAATTAACGTAAGACCGCTGAGGCGCTATACAGTTGCGCTTGCAGAA 4845
 Db 4830 TGCTTGTGATCCATGCTATGATCTCAGAAAGATACAGGCTCAAAAGCTTCAGTCCA 4889
 Oy 4846 GAACAGTTCCGCTATGCTCATCTCTGTTGCGAAGTACAGATACAGGCGTGCAG 4905
 Db 4890 GAACAAATTAATGTGTCTATCTTTCATTCGCGAAGTATTAATATACGTGTGTGAG 4949
 Oy 4906 AAGCTACGTGACCAACAGCTCTGTTTTCAGGCGTGTGACACCGGCTGTACACCC 4965
 Db 4950 AAGATCCAAATCTCCAGCTATATGTTCTCACCGAAAGTGGCTGATATATCATCA 5009
 Oy 4966 AGGAAGTA 4973
 Db 5010 AGGAAGTA 5017

RESULT 5
 us-09-991-258-2
 ; Sequence 2, Application US/09991258
 ; Patent No. US20020141975A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Olmsted, Robert

APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Caley, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swenstrom, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: 01113.0001U3
CURRENT APPLICATION NUMBER: US/09/991.258
PRIORITY FILING DATE: 2001-11-16
PRIORITY APPLICATION NUMBER: 09/902,537
PRIORITY FILING DATE: 2001-07-09
PRIORITY APPLICATION NUMBER: 60/216,995
PRIORITY FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 7479
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975a1e =
NAME/KEY: CDS
LOCATION: (1)...(7479)
US-09-991-258-2

Query Match 42.2%; Score 2111; DB 10; Length 7479;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3209; Conservative 0; Mismatches 1740; Indels 24; Gaps 3;

QY 25 ATGAAAGATTACGTTGACCTTGAATGCTGACGAGCCGCTGATGCTTACAGCGG 84
DB 1 ATGAGAGAAAGTTCACTGTTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 85 AGCTTCCAGAAATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
DB 61 AGCTTCCAGAGTTTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 145 AGAGCGTTTTCGATGTTGCAACAAAGCTCAATTGAGAGAGAGAGAGAGAGAGAG 204
DB 121 AGAGCGTTTTCGATGTTGCAACAAAGCTCAATTGAGAGAGAGAGAGAGAGAGAG 180
QY 205 ATCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 264
DB 181 ATCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 265 ATCTGCGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 324
DB 241 ATCTGCGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 325 AAGAAA-----AGTGAATTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 378
DB 301 AAGAAAAGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 379 GAAATCATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 438
DB 361 GAAATCATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 439 TGTAGGATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 498
DB 421 TGTAGGATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 499 TCAATCTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
DB 481 AGTCTATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 559 ACCCTTTTATGACAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618
DB 541 ACCCTTTTATGACAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

QY 619 GAGAGAGAGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678
DB 601 GAGAGAGAGAGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 679 AGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
DB 661 TCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 739 TTCTCGGTTGTTCAAAATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
DB 721 TTCTCGGTTGTTCAAAATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 799 CAAAGCTGTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 858
DB 781 CCGCTGTATTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 859 AGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 918
DB 841 AGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 919 GAGAGCTTGGCTTCACAAATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 978
DB 901 TCAAGCTATGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 979 CTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038
DB 961 TTAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1039 CAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
DB 1021 CAAATGACTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1099 GGGCTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1158
DB 1081 GGGCTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1159 AACTATCTATTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1218
DB 1141 AACTATCTATTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1219 GACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1278
DB 1201 GACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1279 TGGGCTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1338
DB 1261 TGGGCTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1339 AAGAAATGACCTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1398
DB 1321 AAGAAATGACCTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1399 GATATGAGCTTCCGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1458
DB 1381 GATATGAGCTTCCGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1459 GCTATTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1518
DB 1441 GCTATTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1519 GCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1578
DB 1501 GCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1579 ACCGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1638
DB 1561 ACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1639 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1698
DB 1621 CCTGTGTGCTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1699 ATACTTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1758

Db 1681 GTGCTTTCCTCCGACGGCTGTACTCAAGAGTGAATAATATCTTGATCATCCCTCTCGCT 1740
Qy 1759 GAAACAAGTCTGTATATGACTACAAAGGTAGGGAGGAGATACAAAGTGAAGCATAC 1818
Db 1741 GAAACAAGTCTGTATATGACTACAAAGGTAGGGAGGAGATACAAAGTGAAGCATAC 1800
Qy 1819 CACGGTAAAGTCTATGTATCAGAAAGGAGCGGCTGCTTCAAGACTTCCAGGCTTGG 1878
Db 1801 CATGGTAAAGTATAGTGTGCTCAGAGGAGATGCAATACCGTTCAGAGACTTTCAGCTCTG 1860
Qy 1879 AGTGAAGCGCTACGATCGTTTCAAGAGAGGAGTTCTTAAACAGATACCTGACACAC 1938
Db 1861 AGTGAAGTGCACCATTTGTGTACAAAGAGCGTAGTTCTGTAAACAGGTACCTGACCAT 1920
Qy 1939 ATCGCAATCAACGAGAGGAGCGCTTAAACATGACAGAGTACTATTAAGTCTTAAAGCT 1998
Db 1921 ATTCGCAACATGAGGAGCGCTGAAACATGATGAAGATATTAACAAACCTGTCAAGCC 1980
Qy 1999 CAGGACACAGACTCAGATACGTTCTTGATATGACGACGAAAGTGTGTTAAGGAGAA 2058
Db 1981 AGCGAGACAGAGCGGGAATACCTGTACACATCGACAGAAACAGTGTCTCAAGAAAGA 2040
Qy 2059 GACGCAAGTCCCTTGTGCTTACCGGTATCTGTATGATCCACATTTCAAGATTGCG 2118
Db 2041 CTAGTCACTGGGCTTAGGCTCAGACGAGCGAGTGTGATTCCTCCCTTCATGAATTCGCC 2100
Qy 2119 TACGAGATCTCAGACACGACAGACACCTCCAAAGTCCCAACCATCGAGTCTAT 2178
Db 2101 TACGAGATCTCAGACACGACAGACCGCTCTTACCAAGTACCAACCATAGGGGTGTAT 2160
Qy 2179 GAGAGTCCAGGTTGAGTAAATCTGGATTCATCAAAACGCTGTGACTTAAGAGATCTG 2238
Db 2161 GGCGTCCAGATCAGGCAAGTGTGCACTTAAAGAGGAGTCAACAAAGAAATCTA 2220
Qy 2239 GTTGTAGTGCAGAAAGAAACTGCGCAGAAATCATCAGGAGTAAAGAGATGAGA 2298
Db 2221 GTGTGAGCGCCAGAAAGAAACTGTGCAAAATTAAGGAGCGTCAAGAAATGAAA 2280
Qy 2289 CGTATGATGTTGCTGCTAGACCTGTGCATTCAGTCTTCTAAATGGGTTAAGACCCC 2358
Db 2281 GGGGTGAGCGCTCAAGTGTGAGACCTGTGACTCTTGAAGTGAAGTGAACACCCC 2340
Qy 2359 GTTAACTCTGTATCATTTGATGAGGATTTGCTGCTGAGCAGAGGAGCGTGTGCACTG 2418
Db 2341 GTAGAGACCTGTATATGAGAGACCTTTGCTGTCTGACAGGTACTCTCAAGACGCTC 2400
Qy 2419 ATTGCATCTGTCAAACTTAAGAAAGTGTATTTGTCGGGGAGCCCAAAACAAATGCGCTTC 2478
Db 2401 ATAGCATTAATAGACTTAAGAAAGGCACTGTGCTGCGGGAGTCCAAACAGTGGCTTTT 2460
Qy 2479 TTATAAGATGATGCTGTAAAGTATCTTTAACATGACATATGACATGAAGTACAT 2538
Db 2461 TTATAAGATGATGCTGTAAAGTATCTTTAACATGACATATGACATGAAGTACAT 2520
Qy 2539 AAAAGCATCTTAGAGGAGTGAACAGACTGTAACCCGATCTGTCTCAACCTCTTCTAC 2598
Db 2531 AAAAGCATCTCTCGCGTTGCACTAAATCTGTGATTTGGTGTCTCAACCTTGTTTAC 2580
Qy 2599 GACAAAGCATGAAAGCGGTTAAACCATGTGCTGTAATAATCATATGATATACACAGG 2658
Db 2581 GACAAAGCATGAAAGCGGTTAAACCATGTGCTGTAATAATGATTTGATGACATACGCG 2640
Qy 2659 AACCAAAAGCGGCAAAAGATGATCTGATCTTAACTGTCTTCAAGAGTGGGTGAACAG 2718
Db 2641 AGTACCAAAAGCGGCAAAAGATGATCTGATCTTCACTGTCTTCAAGAGTGGGTGAACAG 2700
Qy 2719 CTACAGATTTGCTCAAAATCAAGAAATCATGATGAGGCTGCAATGCCAAGACTTACG 2778
Db 2701 TTGCAAAATGATTAAGAAAGCAAGAAATTAAGAGGAGCGCTCTCAAGAGGCTGAC 2760
Qy 2779 CGAAAGGCGTTTATGCTGTGATGATCAAGTCAACGAGAAATCCACTTACTGCGAGCT 2838

Db 2761 CGTAAAGTGTATGCTCGGTTCCGATACAGAGTGAATGAATAATCTCTGTACACCCACC 2820
Qy 2839 TCTGAGCAAGTGAAGCTCTTACTTACACGCAACAGAAAGCCATTTGCTGAAAGCGCTA 2898
Db 2821 TCGAAGCAATGTAAGCTCTTACTTACACGCAACAGAGCCGATCTGTGTGAAGAACTACTA 2880
Qy 2899 GCTGTGATCTCTGATTAAGACACTTACACTTAAATATCCCGGGATTTTCAAGGCTTCA 2958
Db 2881 GCGGCAAGCCATGATTAAGAAACACTGATCTGCAAGTACCTGGGAATTTTCACTGCCACG 2940
Qy 2959 TTGAACGATGTCGACGCGGAACAGACGCACTTAATGACGCGTCTTGAATAGCGGAG 3018
Db 2941 ATAGAGAGTGGCAAGCAGACATGATGTCATATGAGGCAATCTTGAAGAGACCGAGC 3000
Qy 3019 ACAGCTATGTTGTTCCAGAAATTAAGTGAACCTCTGCTGGGGAAGCTTTAGAGCAATC 3078
Db 3001 CTACCGAGCTCTTCCAGAAATTAAGCAACCTGTGTGGGCAAGGCTTTAGTGTGCGGTG 3060
Qy 3079 TTGGCACAGGCAATGTGCTGACGAGACAGAGTGGAGAGCTTGCACCATTTCAAG 3138
Db 3061 CTGAAGACCGCTGCTGATGATGATGACATGCACTGAACATGAAGAACCTGTGATTTTGA 3120
Qy 3139 CATGACAGAGGTACTCACTGAATGCACTGAATCTTCTTTGCAACAGGTTCTTTGGA 3198
Db 3121 ACAGCAAAAGCTCACTGACAGAGATGATTTGAACCAATATGAGTGTGAGGTTCTTTGGA 3180
Qy 3199 GTAGACTGACAGTGGGTTATTTTCGCTCTTACCGTGTGCACTTACTTAAGGATCAG 3258
Db 3181 CTGATCTGAGCTCGGCTATTTTCTGCAACCCACTGTCTCGTTATCATTAAGAAATAT 3240
Qy 3259 CACTGGGATTACTGCGCAGGAGAAACATGATGAGGCTTATTAAGAGTGAAGAGAG 3318
Db 3241 CACTGGGATTACTGCGCAGGAGAAACATGATGAGGCTTATTAAGAGTGAAGAGAG 3300
Qy 3319 TTGTCAAGGCGATATCCGTGATCAAAACGCGTTGACAGGAGGTAAGTATTA 3378
Db 3301 CTCTCTGAGAGTACCCCAACATGCTGCGGCAAGTTGCACTGGAAGTCTATGACATG 3360
Qy 3379 AGGAATTAATACATCAAGAGTACTCTTCCAACTTAATGTGTCTTAAATGCGCGG 3438
Db 3361 AACACTGTATACCTGCGCAATTAATGATCCGCGATTAACCTTAAGTCACTGAAGAGAA 3420
Qy 3439 TTGCCCCACTGTTGATCTGTGAACCAAAAGACAGGTAACATCAAGCGGATTC 3498
Db 3421 CTGCTCATGCTTTAGTCTCTCACTAATTAACCAAGAGTACTTTTCTTCAATTC 3480
Qy 3499 CTATCTAAGATGAAGGCGAAATCTGTGTGTGATGCGGATCTTATCAGATTCAGAGG 3558
Db 3481 GTCAAGCAAAATGAAGGCGAGAACTGTCTGTGTGTCGGGAAAGTTGTCCGTCCAGGC 3540
Qy 3559 AAGAAATGATGATCAGTGGTCTCATTTGCCCATTAATCAGTGTGATCTGATTTG 3618
Db 3541 AAAATGTTGATGCTGTGTGTCAGACCGGCTGAGGCTACTTCAAGCTCGGCTGATTTA 3600
Qy 3619 GGAATACCTAGCCATGTCGTAATATGACATTTCTTTGCAATGTTAAGAACCCGATC 3678
Db 3601 GGCATCCAGGTGATGTGCCAAATATGACATTAATTTGTTAATGTGAAGAACCCCATAT 3660
Qy 3679 AGGAACATCACTCAACAAAGTGCAGAGATCACCTTATCCACACAGATCTAACGTT 3738
Db 3661 AAATACATCACTCAACAAAGTGCAGAGATCACCTTATAGCTTATGATTTGACAG 3720
Qy 3739 AAGGCTGTCAACCACTGAAACCTGCGGAAACATGTGTGCTATAGGATATGGCTTGGCT 3798
Db 3721 AAAGCTTGTCTGATCTGAATCCCGGCGGAACTGTGTCAACATAGGTTATGTTACCT 3780
Qy 3799 GATGCGCAACCGGAATATATCATCTGCGGTGCAAGCTCATTAAGTTAACCGGTTC 3858
Db 3781 GACAGGCGCAAGCAAGATATGTGTCTATTAACGCGGCACTTCAAGTTTCCCGGTA 3840
Qy 3859 TGTGAGCTTAAGAACACTGCGGAAATATGAGGTTCTTCTGTTCTTCCGCAAGGAC 3918
Db 3841 TGCAACCGAAATCTCACTTGAAGAGACGGAATTTCTGTTTGTATTCATTGGGTACGAT 3900

```

QY 3919 AACGGACACACATGACAGAGACTCGGTGAGTGTGACCAATCTATCAA 3978
DB 3901 CGCAGAGCCGTATGACCAATCTTACAGCTTTCATCACTGACCTGACCAATTTATACA 3960
QY 3979 GGGTCACCAAGTACGAGGACGAGAGCTCCAGCTGACAGAGTATGAGAGTGTGACATT 4038
DB 3961 GGTTCAGATCTCCAGAGCCGAGTGTGACCTCATATCATATGTGTGCGAGGGGATATT 4020
QY 4039 AGCAGAGCCGTGACCAAGCTATGTTATCTGTCTATATGCAAAAGTCAACGAGTTCC 4098
DB 4021 GCCAGGCGCACCGAGAGGATATTAATGCTCTTAACAGCAAGGCAACCTGCGCGA 4080
QY 4099 GGAAGTGGCGGTGACTGTACCGAAAATGCGGCGTCTTGTATGACAGCAATAGCT 4158
DB 4081 GGGGTGTGGAGCGCTGTATTAAGAGTTCCCGAAAAGCTTCGATTTACAGCCCATGAA 4140
QY 4159 GTCGGGACGCTAGACTTGTGACAGCAAGACC--GCTCATCATATACATGCTGTAGACC 4215
DB 4141 GTAGAAAAGCGCAGCTGTGTAAGAGTGCAGCTAAACATATCATTTCAATGCGTGAACA 4200
QY 4216 AATTTTCTAAGATCCGGAACCGAGGCGACCTTAAGCTCGACGCTGCTTACATGAGC 4275
DB 4201 AACTTCAACAAAGTTTGGAGGTGAAGGTGACAAACAGTTGGCAGAGGCTTATGAGTCC 4260
QY 4276 ATAGCGTCATGTCGCAAGCGTGAAGGATTAACAAAATATGATACGCTACGTCAACC 4335
DB 4261 ATGCTAAGATTTGTCACAGATTAACAATTACAGTACGATGAGATTTCCATTTGTCCACC 4320
QY 4336 GGCATCTATTTGTGTGCAAAAGATCGAGTGAATGCAATCATTTGATCACTGTGTCAGCT 4395
DB 4321 GGCATCTTTTCCGGGAACAAAGATGCACTAACCAATCATTTGAAACATTTGCTCAGACT 4380
QY 4396 TTGCACTAAGGATGCGGATGTCACCATATATGCTTGTGATTAACAATGGGAGAACAGG 4455
DB 4381 TTAAACACCACTGATGAGATGATGACATATATGTCAGAGGCAAGAAATGGAAATGACT 4440
QY 4456 ATATCGAGGCGCTTCAACCGCAAAAGCGTCGAAATTTGGA-----T 4500
DB 4441 CTCAGAGAGAGATGCGCTAGAGAGAGAGAGTGAAGAGATATGATATCCGACGACTCT 4500
QY 4501 GATGACAAAGCAGTAGACATTTGCTGAGGGTCCACCCAAACAGCTTTTGGCAGAGC 4560
DB 4501 TCAGTGAACAGAACCTGATGACAGCTGGTGAAGGGTGAATCCGAAGAATTTTGGCTGGA 4560
QY 4561 AGACCAAGTTACTCGCTCATATGAGGCGCAAGTTGATATCTCTGGAAGTACACGATTC 4620
DB 4561 AGGAAGGCTACACACAAAGCATGGCAAACTTTCATATTTGGAAGGACCAAGTTT 4620
QY 4621 CATGACACCGCCAAAGACATTTGCGAAATTCATGCAATGTGCCCAACAAATGAGGCT 4680
DB 4621 CACAGGCGGCGCAAGAGATATGACAGAAATTAATGCAATGTGCCCGTTGCAACGAGGCC 4680
QY 4681 AATGAGCAGATTTTCTGTATCATCTGTGGGAGAGATATGTCACAGATCGCTCCAAATGC 4740
DB 4681 AATGAGCAGATGATCATGTATATCTCGGAAAGACATGACAGATTAAGTTCGAAATGC 4740
QY 4741 CCAATGAGAGATGAGAGCGCTGTGCTCACTACACACTTCCATGCTGTGTATATAC 4800
DB 4741 CCCGTCGAAGAGTGGGAAGCTTCCACACACTGACAGCTGCTGTGTGTGATTCAT 4800
QY 4801 GCTATGACGCTGAGCGGCTATACAGGTTGCGCTTCGGAAGAAAGAACAGTTGCGCGTA 4860
DB 4801 GCCATGACTCCAGAAAGAGTACAGGCTTAAAGCTTACGCTCCAGAAACAATTAATCTGTG 4860
QY 4861 TGCTCATCATTTCTGTGTCGCAAGTACAGATTCACAGGCGTGCAGAGCTACAGTGACG 4920
DB 4861 TGCTCATCTTTTCAATGTCGGAAGTATGAAATCACTGTGTGCAAGAAATCAATGTCTCC 4920
QY 4921 AAACCACTCTGTTTTCAGGCGTGTACACCGGCTGTACACCCGAGGAAGTA 4973
DB 4921 CAGCTATATATTGTTCTCACCGAAAGTGTGCTCGTATATTTATTCACAGGAAGTA 4973

```

```

RESULT 6
US-09-994-412-3
; Sequence 3, Application US/09994412
; Publication No. US2003059943A1
; GENERAL INFORMATION:
; APPLICANT: Landstrom, Kenneth
; TITLE OF INVENTION: Inhibition of Expression of a Target Gene
; FILE REFERENCE: 20787
; CURRENT APPLICATION NUMBER: US/09/994,412
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 10610
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pSFV2gen(PD) vector
US-09-994-412-3

```

```

Query Match 31.3%; Score 1567.2; DB 11; Length 10610;
Best Local Similarity 58.4%; Pred. No. 0;
Matches 2906; Conservative 0; Mismatches 2038; Indels 36; Gaps 8;

```

```

QY 31 AGAATTCAGCTGACTTGAATGCTGACAGCCGATGTCAAGTCTTACAGCGAGCTT 90
DB 96 AAGTGATGTTGATTTATGAGGCTGACAGCCATTCATCAAGTCTTTGCAAGAGCATTT 155
QY 91 CCAAAATTTGAGATCGAAGCAAGAGCTGACATGACATGACATGCAATGCGAGAGCG 150
DB 156 CCGTGTGAGGATGAGGATGATTTGAGAGTCAACCAAAATATCATGCAATGCGAGAGCA 215
QY 151 TTTTGCATGTGCAACAAAGCTCATTTGAGAGGAACTGACCGGACCAAGTTATTTG 210
DB 216 TTTTGCACCTGGCTACCAATTTGATCGACAGAGACTGACAAAGACACATCATCTTG 275
QY 211 GACATTGGAAGTGGCCGCTGACAGCATGCAATTTCAATCAGCGCTTCAATTTGATGTC 270
DB 276 GATATCGGAGTGGCTTCCAGAGAAATGATGTCTAGCAAAATACCACTGGTATGC 335
QY 271 CTTATGTAAGCGCTGAAAGACCGGACAGACTCAACAGTATGACAGAAAGACT-----T 324
DB 336 CTTATGCGCAGCGGAGAAAGACCCGAAAGGCTGATGCTACGCAAAAGAAAGAGGCGAGCG 395
QY 325 AAGAAAGTACATTTACCGAACAAATAGCTCTTAAGCGGCGACCTGTGGAATGC 384
DB 396 GCTCCGGAAGGTGTGATAGAGATGCGAGAAATATCAACCGACTGCGAGACCGTC 455
QY 385 ATGTCAACACAGACGAGAGACTCCATCTCTGTATGACACAGCGCGAGGTGAG 444
DB 456 ATGGCTAGCGCAGACGCTGAATCTTCACTTTTGCTGTGATACAGCTGACGATGCT 515
QY 445 TACTTTGGAAGTGAAGATATCAAGATGTGTACGACATGACATGACCGACATATC 504
DB 516 ACGGACCGCAAGTGGCGCTGATACAGAGAGTGTATCTGTATATGACCAACATGCTG 575
QY 505 TACCACGAGCGCTTAAAGAGTGAACAATTTATCTGATAGGCTTTGACAGACCCCT 564
DB 576 TACCATCAGCGATGAAGAGTGTGAAACGCGATGATGATGTGGTGTGACACACCCG 635
QY 565 TTTATGTACAAAACATGAGAGTCTTACCTTACTTACAAACGAACTGGGCTGAGAG 624
DB 636 TTTATGTTTACCGGCTAGAGAGGCGGTATCAACTTACGCAACAAATGCGAGAG 695
QY 625 AGAATATTGAAGACATTAATTTGCTGCTGATACCTAGATCTTTCAGAGAGAGGCTT 684
DB 696 CAGGTGTACAGGCGCAAGAAATATGAGCTGTGTGACACATCTTGACTGAGGGAAGATC 755
QY 685 GAAAACTCTCATCTTTAGAGAAAGAGGCTCCAACTCTAATTAAGATCATTTCTCG 744
DB 756 GGCAAACTGTCAATTTCTCCGCAAGAACAAATTTGAACTTGTGCGACAGCTATGTTCTCG 815

```

Qy	745	GTGTGTTCCAAACAATCTACACGAAAGATGATCACTGTAAAGTACTGSCATCTTCCAAAC	804
Db	816	GTAGAGATTACATTGTACACTGAGACGAAAGACTAAGAGAGCTGGCACTTACCTCC	875
Qy	805	GTGTTCCACTTGAAGGAAAGTCTTAACCTTCAAGGTAGATGGGACCATTGTCAGCTGT	864
Db	876	GTATTTCACCTGAAAGGTAAACAATTCCTTTACTGTAGTCCGATTCATGTCATGT	935
Qy	865	GAAGGTTACCTGCATCAAAAAGATAACGATCAGCCAGACCTATACGTTAAAGTTGAGAAC	924
Db	936	GAAGGGTACGTAGTTAAGAAAATACATATGTGCCCCCGGCTGTACGGTAAACGGTAAAGG	995
Qy	925	TTGGCGTCCACAATGSCATCGGAGGGGTTCTTGTAGTTTGCAAAAGTCACAGATACGCTGGC	984
Db	996	TACCCCGTGAAGTATACACGGGAGGAGATCTCTAGTGTCCAGACACAGACACTGTCAAA	1055
Qy	985	GGCGAAGAGGTTCTTTTGTGTGTGTATCGTATGTACCAGCCACACTTGGCATCAATG	1044
Db	1056	GGAAAGAAAGCTCATTTCTCTGTATGACCTACCTCCCTCAACCATCTGTGATCAATG	1115
Qy	1045	ACAGGGATTCCTGGCAACTGACGTTAGTATGTGATGACGCAAAAACATATGTGTGGGCTC	1104
Db	1116	ACTGCGATCTAGGGAACCGACGTCAACCCGAGAGACGCAAGAACTGTATGTGGAATG	1175
Qy	1105	AACCAAGAGATTTGCTGTCATATGTAGACGCGAAAGAAATACTAAACACATGSCAATCTAT	1164
Db	1176	AATCAGAGGATAGTGTGTGTAACGAAAGAACACAGGAAACACTTAACGATGAAACATAT	1235
Qy	1165	CTATTACCACTGTGTGCGCCAGGCGTTTCCAGGTGGGGCGGTGAACATCGTCCGACTTG	1224
Db	1236	CTGCTTCGATGTGGCGCGTGGCAATTTAGCAAGTGGGGGAGGGAATCCAAAGCGACACTT	1295
Qy	1225	GACACGAGAAAGAACTAGGGGTGCGGGAGCGCACTCTTACTATAGGGCTGTGCTGGGCT	1284
Db	1296	GATGATGAAAAAAGCTGTGGGTGTCCGAGAGAGTCACTTACTTGTCTGTGTTGTGGCA	1355
Qy	1285	TTCAAGACCCAGAAAATCATCCATCTTACAAAGAGCTGTGACCAACATTTAAGAAA	1344
Db	1356	TTTTAAAAAGAGAAAGATGACACACATGTACAAAGAAACAGACACCCAGACAAATGTAGAG	1415
Qy	1345	GTACTCTGCGCTTTGACTCATTTGTGATTCACAGCCCTTACACACACGCGCTGCATATG	1404
Db	1416	GTGCTTTCAGGTTTAACTCGTTGTGTAATCCCGAGCCCTATGTGCTTACAGGCTTCGCATTC	1475
Qy	1405	GCGTTCGCCCTCAGGCTCAAGCTGCTGCTTGAAACCACTGTCAAAACCCGACCGGCTATT	1464
Db	1476	CCAGTCAGATCAGCCATTAAGATGCTTTTGGCCAAAGAACCAAGC---GAGAGTTAATA	1532
Qy	1465	ACAAATGGCCGATGTGGAGCATCTGCGTGGCTTTACGACAAAGAGCTGAAGAGTGGCTGCA	1524
Db	1533	CCTGTTCCTCGACGGGTGCTGACGCCAGGGATGCTAAACAAAGGAGAAAGAGAGTGGAG	1592
Qy	1525	GGCGAAGAGATCAGAGAAAGCCCTGCGACCCCTTGCTCCCTGAAATTAAGAAAAAGAACCGTA	1584
Db	1593	GCCGAGCTGACTTAAGAAAGCCTTACACACCCCTGCTCCCATCGCGCGCGGAGAACGGGA	1652
Qy	1585	GAGGCAAGATGAGACTTCATTATGCAAGAG-----GCAGGAGCAGGTAGCGTGGAGACA	1638
Db	1653	GTCTGTGCGAGTCGACGTTGAAGAACTAGAGATACACGACGAGGTGACGGGGTGTGGAAACA	1712
Qy	1639	CCAGAGAGACATACAGGTTGACAAAGTTACCCAGCGCAAGAAAGATTTGGGTCTTACGCT	1698
Db	1713	CCTGCGACGCGGTTGAAGTACGCGACAGCGCAACGACGATCTACTAGGAAATTAACGTA	1772
Qy	1699	ATACTTTACCCCAAGCGGATTTGAATAGTAAAAAATGCGGCTGTATCCACCAATTGGCG	1758
Db	1773	GTCTGTGCTCCCGCAGACCGTGTCTCAAGAGCTCAAGTTGGCCCCCGGACACCTCTAGCA	1832
Qy	1759	GAAACAAGTCTGTATGATCTCACAAAGAGTAGGGCAGGAGATACAAAGTGTGAGCCATAC	1818
Db	1833	GAGCAGGGAATAATTAACAATAACGAGGAGGCGCGCGGTATACAGGTGACAGATAT	1892

QY	1819	JACGGTAAAGCTCATTTGTAACCAAGAGGAGCGCGGCTCCCTGTTTCAAGCTTCACAGCAATTG	1878
Db	1893	GACGGCAAGGGCTCTACTAACCATTGTAATCGGCCATTCCGGCTCTGAGTTTCAGGCTTTG	1952
QY	1879	AGTAGAGCGCTACGATCGTTTTCACAGAGAGGAGTTTCGTAAACAGATACCTTGACCAAC	1938
Db	1953	AGCGAAGCGCGCATATGCTGTATACAGAAAGGAGTTTCGTCAACAGAAACTATACAT	2012
QY	1939	ATCGCAATCAACGGAGAGCGCTAAACATGACGAGAATGATCTATTAAGCTGTAAAGACT	1998
QY	2013	ATTGCCCCCTTACGGAGCCCTCGCTGAACACCGACGAGAGAACTCAGAAAAGTCAGAGCT	2072
Db	2073	GAAAGAACTGACGGCCGAGTACGTTGTGACGTATATAAAAAATGCTCGTCAAGAGAG	2132
QY	2059	GACCGAGTCCCTTGTGCTTAAACCGAGATCTGGTATATCCACCATTTTCAAGATTTGCG	2118
Db	2133	GAACCGTCGGTTTGTGTGTGTGTGGAGAGCTAACCAACCCCGCTTCATGAATTGGCC	2192
QY	2119	TACGAGAGTCTCAAGACACGACGAGCACTTCACAAAGTCCCAACATCGAGTCTAT	2178
Db	2193	TACGAAGGGCTGAAGATACAGGCCGCTCGCACCATATTAAGACTACAGTATGAGAGTCTTT	2252
QY	2179	GGAATGCCAGTTTAGTAAATCTCGAATCATCAAAAGCGCTGTGACTTAAGAAAGATCTG	2238
Db	2253	GGGGTTCGGGATCAGGCAAGTCTGCTATTAATTAAGAGCCCTCGGACCAACACGATCTG	2312
QY	2239	GTTGTGAGTGGGAAGAAAGAACTGGCGCAAAATCATCAGGAGTGTAAAGAGATGAGA	2298
Db	2313	GTCAACGCGCAAGAAAGAACTGCCAGAAATAGTTAAACGCTGAAGAACACCGC	2372
QY	2299	CGTATGAGATGTGCTGTAGGACTGTGCATTCAGTCTTAAATGGGCTTAAAGCACCC	2358
Db	2373	GGGAAGGGGAACAAGTAGGAAAAACATGACCTCATCTCGCTAAACGGGTGTCTGTGGCC	2432
QY	2359	GTTAACACTCTGTACATTGATGAGGCATTGGCTCCCATGACGAGGACGCTGTGGCACTG	2418
Db	2433	GTGACATCTCATATATGTGACGAGGCTTTGGCTTGCATCCCGGATCTGTCTGGCCCTTA	2492
QY	2419	ATTGSCATTCGCAAACT---AAGAAGTGGTATATGGCGGGGACCCAAAACAATGCGGC	2475
Db	2493	ATTGCTCTGTTAACCTCGGACGAATGGTGTATTATGCGAGAGCCCAAGCAATGCGGA	2552
QY	2476	TTCTTTAACTGATGTGCTGAAGTACATTTTAAACATGACATATGCACTGAAGTGTAC	2535
Db	2553	TTCTTCATATGATGACAGCTTAAGTGAACTTCAACCAACAATCTGCACCTGAAGTATGT	2612
QY	2536	CATPAAAGCATCTCTAGAGGTGACACAGACTGTAAACGCCCATCGTCTCCAGCTCTTC	2595
Db	2613	CATPAAAGTATATCCAGACGTTTGACCGGTCTCAAGTCCAGCCATCGTCTACGTTGCAC	2672
QY	2596	TACGACAGCGAATGAAGAAGCGTTAAACCCATGTGCTGATATAAATCATATGATATCCACA	2655
Db	2673	TACGAGGCAAGATGCGACGACCAACCCGCTGCACAAACCCATATATATGAGACACACA	2732
QY	2656	GGGACCAAAAGCCGCAAAAGATGATCTGATTTCTAATCTGTTTCAGAGATGGGTGAAA	2715
Db	2733	GGACAGACCAAGCCCAAGCAGAGACATGTGTTAACATGCTTCGAGGCTGGGCAAG	2792
QY	2716	CAGCTACAGATTGACTACAAAATTCACGAATATGACTGCGGCTGATCGCAGAGACTT	2775
Db	2793	CAGCTGCAATTGAGACTACCGTGACAGAAATCATGACAGACAGCATCTCAGGGGCTC	2852
QY	2776	ACGCGGAAAGCGCTTATGTCTGTCAAGTACAAAATCAAGAAATCCAATCTACTCGCAG	2853
Db	2853	ACCGCAAAAGGGGTATACGCCGTAAAGGCAAAAGGTGATATAAATCTCTGTATGCCCT	2912
QY	2836	ACTTCTAGGACGTTGAAGTGTATTACTTAACGCAACAAAGAAACGCAATGCTGTGAAGAGC	2895
Db	2913	GCGTCGAGACGCTGAATGTACTGTGACGCGCATGAGGATATAGGCTGTGTGAAAAACG	2972
QY	2896	CTACTGTGATATCCCTGGATTAAGACATTACAGCTAATAATATCCCGGGATTTCACGCT	2955

Db 2973 CTGCGCCGCGATCCCTGGATTAAAGTCTCTATCAAACTTCACAGGGTAACTTTACGCC 3032
Qy 2956 TCATTGAGCACTGGCAGCCGGAACAAGCCATTATGGCAGCGCTTCTGTAAGCCG 3015
Db 3033 ACATTGGAAAGATGGCAAGAAACAGCAAAAATATGAAGTGAATGAAGACCGGCT 3092
Qy 3016 CAGACAGCTGATGTTCTCAAAATTAAGGTGAACGTCTGGGGGAAAGGCTTTAGAGCA 3075
Db 3093 GCGCTTGAGACGGCTTCCAAACAAACGAACTGTGTTGGGGAAGAGCTGGTGGCT 3152
Qy 3076 GTCTTGCCACGGCCAACTATGTCTGACGAGACAGAGTGGAGAC---GTTGACCCA 3132
Db 3153 GTCTTGCACTGCGGAAATCAGATTGACAGACAGAGGTGGAGACATTAATTAACGA 3212
Qy 3133 TTCAAGCATGACAGAGCTATCTACCTGAAAATGGCACTGAACCTTTTGGACCAAGTTC 3192
Db 3213 TTTAAGGAGGACAGAGCTTACTCTCAGTGGTGGCTTGAATGAATTTGACCAAGTAC 3272
Qy 3193 TTTGAGAGACCTGGACAGTGGGTATTTCCGCTCCTACCGTCCGACTTACTTACAG 3252
Db 3273 TATGAGTTGACTGGACAGTGGCTGTTTCTGCCCCGAAAGTGTCCCTGTAATACAG 3332
Qy 3253 GATCAGCACTGGGATTAATCTCGCCAGGAGAAACATGTATGGCTTAATAGAGGTAGCA 3312
Db 3333 AACAAACACTGGGATTAACAGACTGTGGAGATGTATGGATTCAATGCCGAACAGCT 3392
Qy 3313 AAGAGTTGTACAGCGCATTCCTGTCATCAAAAAGCGTTGACACAGGAGGTAGCT 3372
Db 3393 GCCAGCTGGAAAGTACATACCTTCTGAAAGGGGAGTGGCATACGGGCAACAGGCA 3452
Qy 3373 GATTAAGGAATTAATACATCAAGGACTACCTCCAAATTAATATGTGTTCCATTAAT 3432
Db 3453 GTTATCGAGAAAGAAAATCCAAACCGCTTCTGTGCTGCAACAATGTAATTCATACAC 3512
Qy 3433 CGCCGTTGCCCACTCGTTGATCGTTGACCAAAAGGACAGGGTACAACTGATCACAGC 3492
Db 3513 CGCAGGCTGCCGACCGCCCTGGTGGTGAATCAAGAAGTTAAAGCAGTAAAGGTTGAG 3572
Qy 3493 GGATTTCTTATTAAGATGAAGGCAATCTGTGTTGATCGCGATCTTATACGATT 3552
Db 3573 TGGCTGTGCAATTAAGTAAAGGGGTACCACTCTGTGTGATGATGATACACTGGCT 3632
Qy 3553 CCAGGAGAAAGATAGATCATGGGTCCATTGGCCCACTAATAC-----CATCAG 3603
Db 3633 TTGCTCGACCGGACGTCATTGTTGTACCGCTGAATGTCACAGGCGCGATAGGTGC 3692
Qy 3604 TGTGATCTGATTTTGGAAATACCTAGCCATGTGGTAAATATGACATTAATCTTTGTCAT 3663
Db 3693 TACGACCTAATTTAGACTCCCGCTGACGCGGCAAGTTGCACTTGTGTAAC 3752
Qy 3664 GTTAGAGACCCGTAACAGAACATATCAACAAAGTGGAGATCAGCTATCCACCAC 3723
Db 3753 ATTCAACGGAATTCAGAAATCACCACTACAGCAGTGTGTGACACAGCCATGAAGTG 3812
Qy 3724 AGCANGCTAAGTGAAGGTGTCCACCACTGGAACACTGGGGGAACATGTGGCTATA 3783
Db 3813 CAGATGCTGGGGAGATGCGCTACGACTGTAAACCCGGGCGGATCT---TGATGGA 3869
Qy 3784 GGGATGGGCTTGTGATCGCGAACCGAATATCATCTGCGGTGGCAGCGCTATTT 3843
Db 3870 GCTTAAGCATACCGCGATTAATCAGCGAAGCCGTGTTTCTCTTAAGCAAAAGTTC 3929
Qy 3844 AGGTTTACCCGTGTGTGACGCTTAAGAACACTGCGGAAAAATAGAGTCTTCTTCTG 3903
Db 3930 TCGTCTGCAAGAGTGTGCGCCGGAATTGTGTACACGAATATACGAAGTGTCTTCTG 3989
Qy 3904 TTTCTTGGCAAGGCAAGCGCAACACACATGACCGAGCAAGCTGGGTGTAGTCTT 3963
Db 3990 TTTCTCAACTTTTGAACAGGAAAGAGACCTCTTACGTAACACAGATTAATCAAGCTG 4049
Qy 3964 GACAACTCTATCAAGGCTCAACAGGTACGAGGAGGAGAGCTCAAGCTTACAGAGTG 4023

Db 4050 AGTCCGCTGATGCGGAGAAAGCATGACACAGCGCGGTGTGACCATCTACAGATT 4109
Qy 4024 ATCAAGGTACATTAACAGAGCGGTGACCAAGCTATCGTTAATGCTGTAATAGCAA 4083
Db 4110 AAGAGACCAACATAGCCACGTGACAGAACCGCTGTGTATACGAGCTAACGCCGT 4169
Qy 4084 GGTCACAGGTTCCGAGTGTGGGTGACCTGTACCCGAAATGCGCGGTCTTTTGTAT 4143
Db 4170 GGAAGTGTAGGGATGTGGTATGATGAGGCCGTGGCGAAAGAAATGGCCGTTAAG 4229
Qy 4144 AGACAGCAATAGCTGTGGAGCGGTGACCTGTGAACGACCAACG---CTCATATA 4200
Db 4230 GAGACCAACACCAAGTGGGCAATTAACAGTCAATGTGCGGCTGTATCCCGTATC 4289
Qy 4201 CAGCTGTAGAACCCATTTTCTAAGATCCGGAACCGAGGGGCACTTAAGCTGCA 4260
Db 4290 CACGCTGACGCTTAATTTCTGTGCAACACATGAAACGGAAGGGACCGGAATTGGCC 4349
Qy 4261 GCTGCTACATAGACATAGCGTCCATCGTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db 4350 GCTGTCTACCGGGAGAGTGCCTGCCGAAGTAAACAGACTGTACAGACAGCTGACATC 4409
Qy 4321 CCGCTACTGTCAACCGGCATCTATTCTGTGGCAAGATGAGTATCATTCGAT 4380
Db 4410 CCGCTGCTGTCAACAGAGTGTTCAGCGCGGAAAGATAGGTGTGACGAATCCCTCAAC 4469
Qy 4381 CACCTGTTCATGTTTTCGACATGACGATGCCGATGCCATATATGCTTGGATTA 4440
Db 4470 CATCTATTCACAGAAAGAGAGCCAGACGCTGACCTGACATCTACAGACADA 4529
Qy 4441 CAATGGAGACACAGATTAATGAGGCACTTACCGCAAAAGAAAGCTGAAATTTCTGAT 4500
Db 4530 AGTTGGAGAAAGAAATTCAGAGAGCTTGAATGACATGAGACGCTGTGAGTGTCAAT 4589
Qy 4501 GATGACACCCAGTGAATGACTTGTGATGAGGTCCAAACCAACAGCTCTTTGGCAGGC 4560
Db 4590 GATGACGTGAGCTGACCAACAGCTTGTGAGTGTGACCCGCGACACAGCTGTGGGT 4649
Qy 4561 AGACCAAGTATCTCCGTCATGAGGGGCAATGTAATCATACCTGGAAGTACAGATT 4620
Db 4650 CGTAAGGCTACAGTACCACTGACAGGCTCGCTGTGATCTTGAAGGTACGAATTC 4709
Qy 4621 CATCAGACCCCAAGGACATTTGCCGAATCCATGCAATGTGGCCCAACAAATCTGAGCT 4680
Db 4710 AACGAGCTGTATTAATGAGGAGATATGACATGATGTGGCCCAACATGCAAGGCA 4769
Qy 4681 AATGACAGATTTGCTTGTATCATCTTGGGAGAGATGTGTCAGCATCCGCTCCAAATGC 4740
Db 4770 AACCAAGATATGCTATATACGCGCTGGGCAAAACAAATGGAACAATCAGATCCAAATGT 4829
Qy 4741 CCGTAAAGAGTGAAGGGGTGTGCTCCACTCACACATTCATGCTGTGTATTAAC 4800
Db 4830 CCGGTGAACATTCGATTATCAACAACCTCCAGAGCAAGTCCCTGTGCGCTAC 4889
Qy 4801 GCTATGACGCTGAGCCGCTATACAGGTTGCGCTCTCGAAGAAAGAAAGTGTGCGCTA 4860
Db 4890 GCAATGACAGAGAACGATGCGCCGCTTAGGTACACCAAGTAAAGATGTGTT 4949
Qy 4861 TGCTCATCTTCTGTGTCGGAATACAGATCAACAGCGGTGCAAGACTACAGTAC 4920
Db 4950 TGCTCATCTTCTTCCCTCCGAAATACATGTATAGTGGGTGCGAAGTAAAGTGCAG 5009
Qy 4921 AAACCACTCTGTTTACAGCGTGTATCAACCGCTTATACCCCAAGAAAGTACGCGGA 4980
Db 5010 AAGGTTCTCTGTTGACCCGACGATCACTTCACTGTGTAGTGTGCGGAAAGTATGCGCA 5069

RESULT 7
US-05-901-106-1
; Sequence 1, Application US/09901106
; Patent No. US20020151067A1
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik

Liljestrom, Peter
TITLE OF INVENTION: DNA Expression Systems Based on
Alphaviruses
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Birch, Stewart, Kolaesch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/901,106
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/920,281C
FILING DATE: 13-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 828-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-241-1300
TELEFAX: 703-241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11517 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHEITICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Semliki Forest Virus
FEATURE:
NAME/KEY: -
LOCATION: 1..11517
OTHER INFORMATION: /label= genome
/note= "Semliki Forest Virus complete nucleotide
sequence, presented as a cloned DNA sequence; see
Figure 5."
FEATURE:
NAME/KEY: CDS
LOCATION: 87..7379
OTHER INFORMATION: /product= "SFV polyprotein"
FEATURE:
NAME/KEY: CDS
LOCATION: 7421..11179
OTHER INFORMATION: /product= "SFV polyprotein"
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-901-106-1
Query Match 31.3%; Score 1564; DB 10; Length 11517;
Best Local Similarity 58.3%; Pred. No. 0;
Matches 2904; Conservative 0; Mismatches 2040; Indels 36; Gaps 8;

Db 216 TTTTGCACACCTGCTACCAAAATTGATCGAGGAGACTGACAAAGACACTCATCTTG 275
Qy 211 GACATTGGAAGTGGCCCGCTCAGACATGACATTCCATTCACCGCTATCTGTATTCG 270
Db 276 GATATCGGAGTGGCCCTTCCAGGAAATGATGTCTAGCACAATAATACACTGCTATGCG 335
Qy 271 CCTATGATAGGCGCTGAAGACCCCGGACGACTACAAAGGATGACGAAAGACT-----T 324
Db 336 CCTATGCGGAGCGGAGAAAGACCCGAAAGGCTCGATGCTACGCAAAAGAACTGGCAGCG 395
Qy 325 AAGAAAGTGAATTACCGACAAAGAACATAGCCTTAAAGCGCGAGACTGCTGAAAGTC 384
Db 396 GCCTCGGGGAAGTCTGATAGAGATCGAGAGAAATATCACCAGCTCGACACCGTC 455
Qy 385 ATGTAAACACGAGACGAGACATCTCTGTATAGCAGACAGAGCCGCTGTAGG 444
Db 456 ATGGCTACGCGAGAGCGTGAATCTCTTCCCTTGGCATACAGACGTCAGTGTGCT 515
Qy 445 TACTTGAAGTGTAGCAGTATACCAAGATGTGTACGAGTCCATGCAACCGCATCAATC 504
Db 516 ACGGACCGGAAGTGGCCGTATACGAGACGTGTATGCTGTACATGCAACCAATGCTG 575
Qy 505 TACCAACGAGCGCTTAAAGAGTTAGACAAATTACTGATAGCCTTTGACACGACCCCT 564
Db 576 TACATCAGCGGATGAAGTGTGAGAACGGGATGATGATGGGTTTGACACACCCCG 635
Qy 555 TTTATGTACAAAACATGCGAGTTCCTACCTCTTACACACGAACTGGGCTAGCG 624
Db 636 TTTATGTTTGAACGCGCTTACGAGCGCCGTATCCAACTGCGCACAACTGGGCGCCACG 695
Qy 625 AGAGTATTGAAGACGTATACATTGGCCCTCGTACTCAGATCTTCAGAGAGCGGCT 684
Db 636 CAGGTGTACAGGCCAGGAATAGACTGTGTGAGACATCTTGACTGAGGAAGACTC 755
Qy 685 GAAAACTCTCAATCCTTAGAAGAGAGGCTCCAACTACTAATATAGATATATTCG 744
Db 756 GGAACAACTGCTCATCTCCGAAAGAGCAATTGAACCTTGGACACAGTATGTTCTCG 815
Qy 745 GTTGTTAAACAATCTACAGAAATGATGATCACTGTTAGTACTGCTGCATCTTCAAC 804
Db 816 GTAGATCTTACATTTGACCTGAGGAGAAAGCTACTGAGAGCTGCACTTACCTCC 875
Qy 805 GTGTCCACTGGAAGAGAAAGTCACTTCAAGTAGATGAGGACATGTCAGCTGT 864
Db 876 GTATTCCACTGAAAGTAAACAATCTTTACCTGTAGCTGATACATGATCATGT 935
Qy 865 GAAGGTACGTCATCAAAAAGATTAAGATCAGTCCAGGACTATACGTTAAAGTTAGAAC 924
Db 936 GAAGGTACGTTAGTTAAGAAATCACTATGCCCCGCGCTGTACGTTAAACGTTAGG 995
Qy 925 TTGGCTTCAACAATGATCGGAGGGTTTCTTGAAGTTCAAAAGTACAGATACGTCGCC 984
Db 996 TACGCGGTGATACAGCGGAGGATTCCTAGTGTGCAAGACCAAGACACATCTCAAA 1055
Qy 985 GCGCAGAGGGTTCTTTGCTGTGTGATACGATGATACAGGACACTTTGGATGATG 1044
Db 1056 GGAAGAAAGTCAATCTCTGTATGCACTTACGCTCCCTCAACCATCTGTATGATTAATG 1115
Qy 1045 ACAGGATTTGGAACATGACGTTAGTGTGATGACGACAAAATCTATTGGTTGGGCTC 1104
Db 1116 ACTGCACTACGAGGACCGAGTCAACCGGAGAGACGACGAAAGTTGTAGTGGATG 1175
Qy 1105 AACCAAGATTTGCTGCTAATGTGAGACGCAAAAGAAATCTAACCAATGACAGATAT 1164
Db 1176 AATCAGAGATATGTTGTGAACGAAACACAGCAAAACATTAACAGATAGAAATAT 1235
Qy 1165 CTATTACAGTGTGCGCCAGCGTTTCCAGGTGGGCGGTGAACATGTCGCGAGCTG 1224
Db 1236 CTGCTTCGATTTGGCCGCTGCAATTAGCAAGTGGGAGGAAATACAAAGCAGACTT 1295
Qy 1225 GACACGAAAGAACTAGGGGTGCGGAGCGCACTCTACTAATGAGCTGCTGCGGCT 1284

Dh 1296 GATGATGAAAAAAGCTGCGGTGTCAGAGAGAGTCACTTACTGTGCTGCTGGGCA 1355
Qy 1285 TTCAAGACCCGAAAAATCACATCCATCTACAAAGAAAGCTGTAGCGAAACAATTAAAGAA 1344
Db 1356 TTTTAAACGAGAGAAATGACACACATGTACAAAGAAACCAACACCCAGACATATGTAGAG 1415
Qy 1345 GTACCTCCCGCTTTGATCTATTGTGATTCACAGCTTACAGCCAGCGGCTCGATATG 1404
Db 1416 GTGCTTCAAGATTAACTGCTGTCAATCCCGAGCCATATGTGTACAGGCTTGGCAATC 1475
Qy 1405 GGCCTTCGCGGTGAGCTTACAGCTGCTGTGAACCAACTGTCAAAACCGGCAACGGTATT 1464
Db 1476 CCAATCAGATCAACCATTAAGATGCTTTTGGCAAGAAAGACCAAGC---GAGAGTAAATA 1532
Qy 1465 ACAATGGCCGATGTGTGAGCATCTGTGCGCTTACAGCAAGAAAGCTGAAGAGTGGCTGCA 1524
Db 1533 CTGTCTTCGACGGGTGTCAGCCAGGAGTGTGAACCAAGAGAGAAAGAGATGGAG 1592
Qy 1525 GCGGAAAGATCAGAGAAAGCCCTGCAACCTTGTCTCCCTGAATATGAAAAAGAGACCGTA 1584
Db 1593 GCCGAGCTGACTAGAGAAAGCTTACACCCCTCGTCCCATCGCGCCGGGAGACGGGA 1652
Qy 1585 GAGGCAAGATGACCTCTATTATGCAAGAG-----GCAAGAGAGATGAGCTGAGACA 1638
Db 1653 GTCTGCAAGCTGACGTTGAAGAACTAGAGTATCAGCAGAGGTGCGAGGGGTGTGAAAAA 1712
Qy 1639 CCAGAGAGACATCAGGGGTGACAAAGTTAACCCAGCGCAAGAAAGATTGGGTCTTAGCCT 1698
Db 1713 CCTGCAAGCGGCTTGAAGATGACGCAACGCCGAACGAGTACCTAGGAAATTAGCTA 1772
Qy 1699 ATACTTTCACCCCAAGCGGCTTGAATATGTAAGAAAACTGCGGTATTCACCCATTGGCG 1758
Db 1773 GTTGTGTCCTCCGAGACCGTGTCTCAAGAGCTTCAAGTTGGCCCCCGGTGACCTCTAGCA 1832
Qy 1759 GAACAACTACTGTATGACTCACAAAGTAGGGCAGGAGATCAAAAGTGCAGCCATAC 1818
Db 1833 GAGCAGGTGAAAAATTAATACATATACGAGAGGCCGCGGTACAGAGTGAAGGATAT 1892
Qy 1819 CAGCGTAAGCTATTGTATCCGAAGAGGACCGCGGTCTCTGTTCAAAGACTTCCAGGCAATG 1878
Db 1893 GACGGCAGGGGTCTTACTACCACTGTGANTGGCCATTCGGTCTCTGAGTTTCAAGCTTTG 1952
Qy 1879 AGTGAAGCGCTGATCGATGTTTTCAACGAGAGGAGTTCTGAACAGATCTCCACAC 1938
Db 1953 AGCGAGAGCGGCACTATGTGTACACGAAAGGAGTTCTGTCAACAGAAATATACCAT 2012
Qy 1939 ATTCGAATCAACGAGAGAGCGCTAAACATGACGAAGATATCTAATAGACTGTAAAGACT 1998
Db 2013 ATTGCGGTTCAAGCAGCGTGTGTAACACGAGAGAGAACTACGAAAGTCAAGAGCT 2072
Qy 1999 CAGGACACAGACTGAGATAGTCTTGATATGACGACGAAAGTGTGTTAAGCGAGAA 2058
Db 2073 GAAGAACTGACGCGAGTACGTGTTCGACGTAAATTAATAATCTGTGTAAGAGAGAG 2132
Qy 2059 GACGCAAGTCCCTTGTGCTTAACCGGTGATCTGTAGATCCACCATTTCAAGATTTGCG 2118
Db 2133 GAACGCTCGGTTTGTTGTGTGTGTGAGAGACTAACCAACCCCGTTTCAAGATTTGCC 2192
Qy 2119 TAGCAGAGTCTCAAGACACGACAGCAGCCTTCAAAAGTCCAAACCATGAGTCTAT 2178
Db 2193 TACCAAGGGCTGAAGATCAAGCGCTCGGCACTATTAAGATAGTAGAGTCTTT 2252
Qy 2179 GGAATGCAAGTTGATAGTAAATCTGGAATCATCAAAAGCGGTGATCTAATAAAGATCTG 2228
Db 2253 GGGGTTCCGGGATAGGCAATCTGTATTATTAAGACCTGTGACCAAAACAGATCTG 2312
Qy 2239 GTTGTGATGCGAAGAAAGAAATCTGCGCAAGAAATCATCAGAGATGTAAAGAGATGAGA 2298
Db 2313 GTCAACAGCGGCAAGAGAGAACTGCCAGAAATAGTTAACGACGTGAAACACACCCG 2372
Qy 2299 CGATAGATGTTGTGCTGAGAGCTGTGATTCAGTGTCTTAAATGGGGTTAAGCACC 2358
Db 2373 GGGAAAGGAGCAAGTAAAGGAAACAGTGAATCTCTGTAACGGGTTGTCTGTGCC 2432

Qy 2359 GTTAACACTGTATACATTATGAGGCAATTTGCCGCAATGCAAGGACGCTGTGCACTG 2418
Db 2433 GTGACATCTCTATATGTGTGAGAGAGCTTTGCTTGCCATTCCGATCTGTGTGCTTGA 2492
Qy 2419 ATTGCCATCTCAAACT---AGAAAGTGTATGTGTGCGGAGACCCAAACAAATGCGGC 2475
Db 2493 ATTGCTCTGTAAACCTCGAGCAAAAGTGTGTATGCGAGACCCCAAGCAATGCGGA 2552
Qy 2476 TTCTTTAATCATGTATGCTCTGAAAGTACATTTTAACATATACATATGACTGAAGTAC 2535
Db 2553 TTCTTCATATATGATGACGCTTAAAGTGAACCTTCAACCAACATCTGCACTGAAGTGT 2612
Qy 2536 CATTAAGACATCTCTAGAGGTGACACACAGCTGTAAACGCGCATGCTCCAGCCTCTTC 2595
Db 2613 CATTAAGATATATCAGACGTTGACGCTTCAGCTACGCGCATGTGTCTACGTTGCAC 2672
Qy 2596 TACGACAAAGCAATGAAGAGGTTAAACCATGTGCTGATTAATAATCATATATATACCA 2655
Db 2673 TAGGAGGCAAGATGGCAGACCAACCGTGCAACAAACCAATATCATATAGACACACA 2732
Qy 2656 GGAACCAAAAGCCGCAAAAGATGATCTGATTTCTAACCTGTTTCAAGAGATGGGTAAA 2715
Db 2733 GAAACGACAAAGCCCAAGCCAGAGAGATGTGTTAACATGCTTCCAGGCTGGCAAG 2792
Qy 2716 CAGCTACAGATTGACTCAAAATAATCAGAAATCATGACTCGGCTGATGCAAGGACTT 2775
Db 2793 CAGCTGAGTTGAGACTACCGTGACAGCAAGTCAATGACAGCAGATCTCAGGCTTC 2852
Qy 2776 ACCGGAAGAGCGTTATGTGTCAAGTACAAAGTCAACGAGATCTACTGTGACG 2835
Db 2853 ACCCGAAAGGGGATATCGCGCTAAGGCAAGATGTAATAATCCTGTATGCCCCCT 2912
Qy 2836 ACTTCTGAGCACTGAAAGTGTACTTATACGCAACGACAGAAAGATGTTCTGGAAGACG 2895
Db 2913 GCGTGGAGCAGCTGATGTACTGTGACGCGACGTGAGATAGCGTGTGTGAAAAACG 2972
Qy 2896 CTAGCTGTGATCCCTGATTAAGAAAGACCTTACAGCTTAATATCCCGGGATTTCAAGGCT 2955
Db 2973 CTGGCCGCAATCTCTGATTAAGAGTCTTATCAAACTTCCACAGGTAATCTTACGCGC 3032
Qy 2956 TCAATTGACACTGTGACGCGCAACAGACGCACTTATGCGACGCTTCTGTATAGCCG 3015
Db 3033 ACATGGAAGAAATGGCAAGAAAGAACAGCAAAATATGAAGTATGAAGACCGGCT 3092
Qy 3016 CAGACAGCTATGTGTTCCAGAAATAGGTGAAGTCTGTGCGCGAAGGCTTATAGACCA 3075
Db 3093 GCGCTGTGACCGCTTCCAGAACAAAGCAAGCTGTGTGGCGCAAAACCTGTGTGCT 3152
Qy 3076 GTCTTGSCACGGGCAACATTTGTGTGACGAGACAGCAGTGGGAGAC---GTTGACCCA 3132
Db 3153 GTCTGGAACCTGCCGAATCAGATTGACAGAGAGATGAGACATATATTAACGCA 3212
Qy 3133 TTCAAGCATGACAGAGGTACTCACTGAATAGCACTGAATCTTTTGCACCAAGTTTC 3192
Db 3213 TTTAAGGAGACAGAGCTTACTCCAGTGTGCTTGAATGAATAATTTSCAACCAAGTAC 3272
Qy 3193 TTTGAGTAACTCTGACATGTGGTTATTTTCCGTTCTTACCGTGCACCTTAATACAGG 3252
Db 3273 TATGAGATTAACTGACAGATGTGCTGTTTTCGCCGGAAGTGTCTCTGTATTAAGAG 3332
Qy 3253 GATCAGACTGGGATTAATCTGCCAGGAAAGAAATGATGTGGGCTTAAATAGAGAGTGA 3312
Db 3333 AACCACTCTGGATTAACAGACTGTGTGGAAGAGATGTATGATTCATATGCCGAACGCT 3392
Qy 3313 AAGAGTTGTACGCGCATATCTGTGATCAAAAAGCGGTGAACAGCAGGAGTATGCT 3372
Db 3393 GCAAGCTGGAAGCTAGACATCTTCTGAAGGGGACATGCGCATGCGGCAAGCAGGCA 3452
Qy 3373 GATTAAGAAATATACATCAAGAGACTACTTCCCAACATTAATATGTGTGTTCAATTAAT 3432
Db 3453 GTTATCGCAAGAAAGAAATCAACCGCTTGTGTGTGTGACATATTAATTTCTATCAAC 3512

OY	3433	CGCGGGTGTGCCCACTCGTGTATCGTTAGACCAAGAACAGAGGACAGGGTACAACTGATCACAGC	3492
Db	3513	CGCAGGCTCCGCGACGCCCTGGTGGCTGAGTACAGACGGTTAAAGGACGTAAGGTTGAG	3572
OY	3493	GGATTCTATCTAAGATGAAAGGACAAATCTGTGTGTGATCGCGCATCTATCAGCAAT	3552
Db	3573	TGGCTGTCAATTAAGTAAAGAGGGTACACACGTCCTGTGTGATGATGATCAACCTGGCT	3612
OY	3553	CCAGGGAAGAAAGTGAAGTCCATGGGTGCATTGGCCCAATAC-----CATCAGG	3603
Db	3633	TTGCCTGCAGCGAGGGTCACTTGGTGTTCACCGCTGAATGTCAACGGGCGCGATAGGTGC	3692
OY	3604	TGTGATCTCGATTTTGGGAATACCTAGCCATGTCCGTAATATATGACATTAATCTTGTGCAT	3665
Db	3693	TACGACTTAAGTTTGAAGCTGCGGCTACCGCCGACAGTTCGACTTGGTCTTGTGTGAC	3752
OY	3664	GTTAAGAACCCCGTACAGAAACCATCACTACCAACAGTCCGAGATCAAGCTATCCACAC	3723
Db	3753	ATTCAACAGGAATTCAGAAATCCACCACTACAGAGTGTGTGCACACCGCATGAAGCTG	3812
OY	3724	AGCATGCTAAAGGTGAAGGCTGTCCACCACTGAACACTGGCGGAACATGTGTGCTATA	3783
Db	3813	CAGATGTCTGGGGGAGATGGCTACGACTGCTTAAACCCGGCGGCATCT--TGATGAGA	3865
OY	3784	GGGATAGGGCTTGGCTGATCGCGCAACCGAAGATATCATCACTGGCGTGGACGCTCATTT	3843
Db	3870	GCTTACGGATACGGCCGATTAATAATCCAGACGCCGTGTGTTCTCTTAAAGCAAGAAATTC	3929
OY	3844	AGGTTTACCCCTGTCTGTCAAGCTTAAGAACCTGCGGAAATATCTAGAGGTTCTCTTGGTG	3903
Db	3930	TGCTGTCAAGAGTGTTCGCCCGGATTTGTGTCAACAGCAATACAGAAGTGTCTTGGTG	3989
OY	3904	TTCTTCCGCAAGACACGCGCAACCAACACATGACGAGACAGACTGGTGTGTGTCTT	3963
Db	3990	TTCTTCCAATTGTGACAAAGGAAAGAGACCTCTACGTAACACAGATGAATACCAAGCTG	4049
OY	3964	GACAAACATCTATCAAGGGTCAACCAAGTACGAGAGGAGGAGAGCTCCAGCCTACAGAGTG	4023
Db	4050	AGTGCCGTGTATGCCGGAGAGACCATGACACCGCCGGGTGTGCACCATCTCAAGAGTTT	4109
OY	4024	ATCAGAGTGAATTTAGCAGAAAGCGCTACCAAGCTATCTGTTAATGTCTAATAGCAAA	4083
Db	4110	AAGAGAGAGACATATGACCAGTGCACAGAACCGCTGTGTGTTAACGACGCTACGCCCCGT	4169
OY	4084	GGTCAACACGGTTCGCGAGTGTGCGGTGCACTGTACCGAAATGCGCCGCTGCTTTGAT	4143
Db	4170	GGAACGTGAAGGGATGGCGTATGACGGGCCGTGGGGAAGAAATGGCCGTACGCCTTAAG	4229
OY	4144	AGACAGCCAAATAGCTGTGCGGACGCTTGAATTTGTGAAGCAACGAACG--CTATCATATA	4200
Db	4230	GGACACAGAACACAGATGGGGCAATTTAAACAGTCAATGTCCGGCTGTACCCCGTATCC	4289
OY	4201	CATGCTGTAGAGACCCCAATTTTTCTTAAGTAGCGGAAACCGGAGGGCGACCTTAACCTCGCA	4260
Db	4290	CACGCTGAGCGCTTAATTTCTTGCACGACTGAAGGGAAGGGAGACCGGAATTTGGCC	4349
OY	4261	GCTGCTACATAGACATAGCGTCCATCGTCAACGCTGACGGATTACAAAATAATTCAGTA	4320
Db	4350	GCTGTCTACCGGGGAGTGGCCGCGGAGTAACAGACTGTACTGACGAGGTAGCCATCC	4409
OY	4321	CCGCTACTGTCAACCGGCATCTAATTTCTGTGGCAAGATGCAATGATGCAATCAATTGCAT	4380
Db	4410	CCGCTGTCTGTCAACAGGATGTGTTACGCGCGGAAGAGATAGCTGCAGCAATCCCTCAAC	4469
OY	4381	CACCTGTTCATGCTGTTTGCACCTAAGGATGCCGATGTCAACATATATTTGGATTGATAAA	4440
Db	4470	CATCTATTCAACAGCAATGAGCGCCACGACGCTGACGTGACCATCTACTGACGAGACAAA	4529
OY	4441	CAATGGGAGACAGAGATTAATGAGGCCATTCAACCGCAAAAGAAAGCGTTCGAATTTCTGGAT	4500
Db	4530	AGTTGGGAGAGAAATATCAAGAAAGCAATTGACATAGGAGACGGCGTGTGAGATTCCTCAAT	4589
OY	4501	GATGACAAAGCCAGTATGACATTGAGCTTGGTCAAGGCTCACCCAAACAGCTTTTGGACAGCC	4560

Db	4550	GATACCTGAGCTGACCA	CAGACTTGTGTGAGTGA	CAACCCGACACAGCTGTGGGT	4649
Qy	4561	AGACAGATTACTCCGTCA	TGAGGGCAAGTGTATT	CATACCTGGAAGGTACGATTC	4620
Db	4650	CGTAAGGGCTACAGTAC	CACTGACGGGTCCGTACT	CGTACTTGAAGGTACGAATTTC	4709
Qy	4621	CATAGACCCGCAAGAC	ATTGCCGAATCCATGCA	ATTGGCCCAACAAATCTGAGCT	4680
Db	4710	AACCAAGCTGTATTGTA	TATGGCAGAGATCTGA	CGTTGTGGCCCAACTGCAAGAGCA	4769
Qy	4681	AATGAGAGATTGCTGT	ATCATCTGGGGGAGATG	TGCCAGCAATCCGCTCCAAATGC	4740
Db	4770	AACGAAACGATATCTTA	TACGCGCTGGGGAACAA	TGACAAATCAATCCAAATGT	4829
Qy	4741	CCAGTAAAGAGTGAGAG	CGCTGTGCTCCAACTCA	CACACTTCATCCTGTGTAATTAC	4800
Db	4830	CCGGTGAACGATTCGAT	TCAATCAACACTTCCAG	ACAGATGCGCTGTGTGCGGCTAC	4889
Qy	4801	GCTATGACCGCTGAGCG	CGTATACAGGTTCGCT	CTTGCGAAGAAAGACATTCGCCGTA	4860
Db	4890	GCAATGACAGAGAA	CGGATCGCCGCTTAGGT	CACACCAAGTTAAAGCATGATGGTT	4949
Qy	4861	TGCTCATCATTCCTGT	TGCCGAAGTACAGATCA	CAGGCGTGCAGAAAGCTACAGTCAGC	4920
Db	4950	TGCTCATCTTTTCCCT	CTCCGAAATACATGAT	GTAGTGGGTGCGAAGGATTAAGTGGAG	5009
Qy	4921	AAACCAAGTCTGTTTTC	AGGCGTGATACCA	CCGGCTTACACCCACAGAAAGTACGGGAA	4980
Db	5010	AAAGTTCTCTCTGTTCA	CCGACGGTAACCTTCA	GTGTGTTAGTCCGCGAAGATATGCCCA	5069

```

RESULT 8
US-09-190-246-1
: Sequence 1, Application US/09190246
: Publication No. US2003018025A1
: GENERAL INFORMATION:
: APPLICANT: Parrington, Mark
: APPLICANT: Li, Xiaomao
: TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYOXINUS VACCINES
: FILE REFERENCE: Parrington et al
: CURRENT APPLICATION NUMBER: US/09/190,246
: CURRENT FILING DATE: 1998-11-13
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 15538
: TYPE: DNA
: ORGANISM: Semliki Forest virus
US-09-190-246-1

```

Query Match	31.3%	Score 1564;	DB 12;	Length 15538;
Beat Local Similarity	58.3%	Pred. NO. 0;		
Matches 2904; Conservative	0;	Mismatches 2040;	Indels 36;	Gaps 8;

Oy	31	AGAAATTACGCTGTA	CTTAGATGCTGACAGCCCGATATGCAAGTGGTTA	CACGGACGCTT	90
Db	2642	AAATGTCATGTTGATTA	TTAGAGCTGACAGCCCAATTCATCAAGTCTTTTGACAAAGCATT	2701	
Oy	91	CCACAATTTGAGATGAG	CAGCAGGCACTGACATGACATGCATGCCCAATGCCAGAGCG	150	
Db	2702	CCGTCGTTTCGAGGTTGAG	TGATCTTTCAGCTGACACCAATATGACATGCAAAATGCCAGAGCA	2765	
Oy	151	TTTTTCGATGTGGCA	CAAAGCTCATTTAGAGCGGAAGTCCAGCCGGACCAAGTTATCTTG	210	
Db	2762	TTTTTCGACCTGGGTAC	CAAAATGATGAGAGGAGACTGCAAAAGCACACTATCTTG	2821	
Oy	211	GACATTTGGAAGTGGCCG	CTCAGACATGACATTTCCATATCCCGCTATCATTTATCTGC	270	
Db	2822	GATATGGCAGTGGCCCTT	CCAGAGAAATGATGTCTACGCAAAATACCACTGCTGATGC	2881	
Oy	271	CTATATGATTAAGCGCT	GAGAAGCCGGACAGACTACAAACGGTATGACGAAGAATT-----T	324	

Db 2882 CCTATGCGCAGCGCAGAAAGACCCGAAAGGCTCGATAGCTACGCAAAAGAACTGCGACGG 2941
Qy 325 AAGAAAAGTGACATTACCGACAAAGAACTATACCTCTTAAGCGCGACACTCTCTGAAATC 384
Db 2942 GCCTCCGGGAAAGGTGCTGATAGAGATCCGAGAAAATCACCGACTCTGCAACCGCTC 3001
Qy 385 ATGTCAACACGACGACGAGACTCCATCTGTGTATGACACAGACGCGACGCTGATAG 444
Db 3002 ATGGCTACGCGACGCTGAATCTCTTACCTTTGCTGACATACGACGTACGCTGTCT 3061
Qy 445 TACTTGGAGTGTAGCAGTATACCAAGATGTATACGAGTCCATGCAACGCACTCAATC 504
Db 3062 ACGGACCGCAAGTGGCGGTATACAGGACGATATGCTATCAATGACCAACATCGCTG 3121
Qy 505 TACCAACAGCGGCTTAAAGAGTTAGCAATTTACTGAGTGGCTTGAACAACCCCT 564
Db 3122 TACCATCAGCGGATGAAAGGTGTCAAGACGCGTATGGATTGGGTTTGAACACCCCG 3181
Qy 565 TTTATGTACAAACATGCGCAGGTTCTTACCTACTTACACAGAACTGGGCTGACGAG 624
Db 3182 TTTATGTTGACGGCTAGCAGCGCGGTATCAACCTACGCAAACTGGGCGAGCGAG 3241
Qy 625 AGAGTATGGAAGCAGTAACTTGGCTCGTAACTCAGATCTTCAGAGAGAGGCTT 684
Db 3242 CAGGTGTACAGGCGAGGACATAGGACTGTGTGAGCATCTTGAACGAGGAAAGCTC 3301
Qy 685 GGAAGAACTCTCAATCTTGAAGAGAGGCTCCAACTTAAATAGATCATTTCTCG 744
Db 3302 GCGAAACGTCTCATCTCCGAGAAAGCAATTGAACCTTCGACACAGTATGTCTCG 3361
Qy 745 GTTGGTTCAACATCTACAGAAAGATAGATCACTGTACGTAGTGCGCATCTTCAAC 804
Db 3362 GTAGATTTAATTTGTACTGACGAGCAAGAAAGCTACTGAGAGCTGCACTTACCTCC 3421
Qy 805 GTGTCCACTTGAAGAGAAAGTCTTAACTTCAAGGTAGATGTGGAGCCATTTGCACTGT 864
Db 3422 GTATTCGACCTGAAAGGTAAACAATCTTTACCTGTAGTGCGATACCATGTATCATGT 3481
Qy 865 GAAAGGTACTCTATCAAAAAATACATCAAGCCAGCACTATACGTTAAAGTTGAGAAC 924
Db 3482 GAAAGGTACTGTATTAAGAAATCACTATGTGCCCCGCTGTACGTTAAACGTTAGGG 3541
Qy 925 TTGGGTCCACAATGCGATCGGAGGGTTCTTGAATGTCAAAGTACAGATACGCTGGC 984
Db 3542 TACGCGTACGATACGCGGAGGGATTCCTAGTGCAGAGACCAAGACACTGTCAA 3601
Qy 985 GCGCAGAGGGTTCTTTTGTGTGTGTACGTATGACGACACACTTTGCGACATG 1044
Db 3602 GGAGAAAGAGTCTATTCCTGTATGACCTACGTCCCTCAACCATCTGTGATCAATG 3661
Qy 1045 ACAGGAGTCTGCGCACTGACGTATGTGTGATGACGACAAAACTATTTGGTGGGCTC 1104
Db 3662 ACTGGCAATCTAGGACCGACGCTCACCGGAGGCGACAGAAAGTTTGTAGTGGGATG 3721
Qy 1105 AACCAAGAGTTGTGCAATGTGTAGGACGCAAAAGAAATCTAACACAAATAGCAACT 1164
Db 3722 AATCAGAGGATATGTTGTAAAGGAAACACAGCAAACTAACACATAGAGAACTAT 3781
Qy 1165 CTATTAACGATGTGCGCCAGCGCTTTTCAAGTGGGCGCGTGAACATCTGTCCGACTTG 1224
Db 3782 CTGCTTCGATTTGTGGCGGTGCGCATTTAGCAAGTGGGAGGAAATACAGGACGACTT 3841
Qy 1225 GACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTTACTAGTGGCTGTCTGGGCT 1284
Db 3842 GATGTATGAAAACTCTGGGTGTCCGAGAGAGGTCACTTCTGCTGTGGGCGCA 3901
Qy 1285 TTCAAGACCCAGAAATCACTCACTTCAAGAGAGCGGTGTAAGCAAAATTAAGAA 1344
Db 3902 TTTAAAGGAGAGATGACACACATGTACAAAGAAACAGACCCCAACATATGTAG 3961
Qy 1345 GTACCTGCGCTTTGTGACTATTTGTGATTCACGCTTACAGGCAAGCGGCTGATATG 1404

Db 3962 GTGCTTTCAAGATTAACTCGTTCTGATCCGAGCCTATGTCTACAGGCTCTGCAATC 4021
Qy 1405 GGTCTCCCGGTATGGCTCAGCTGCTGTGAACCAACTGTCAAAACCGGACCGGCTATT 1464
Db 4022 CCACTCAGATCAGCATTTAAGATCTTTTGGCCAAAGAACCAAGC--GAGATTATTA 4078
Qy 1465 ACAATGCGCATGTGAGCATCTGCTGTCTTACAGCAAGAACTGAAAGATGTGCTCA 1524
Db 4079 CTGTCTCGACGCTGTGACGCAAGGATGTGAAACAAAGAGAAAGAGATGTGAG 4138
Qy 1525 GCGGAAGATCAGAGAAAGCTTGCCACCTTGCTCTCTTAAATAGAAAAAGAACCTTA 1584
Db 4139 GCGGAGCTGACTAGAGAAAGCTTACCACTCGTCTCCCTCATCGCGCGGAGACCGGA 4198
Qy 1585 GAGCGAATATAGACCTTATAGCAAG-----GCAAGAGAGGTAGCTGAGACA 1638
Db 4199 GTCTTCAGCTGACGTTGAAGAACTAGATATCACCAAGTGCAGGGGTGTGGAACA 4258
Qy 1639 CCAAGAGACATACAGGAGTGAACAATTACCAAGCGCAAGAGAAATTTGGTCTTACGCT 1698
Db 4259 CTTCGACGCGCTTGAAGATCACCGCACAGCCAGACGTAATCTACTAGAAATTAAGTA 4318
Qy 1699 ATACTTTCACTCCAGCGGATTAATAGTAAACTGAGCTGTATCCACCAATTTGGC 1758
Db 4319 GTTCTGTCCCGAGACCGGTCTCAAGAGCTCCAAAGTTGCCCCCGTGCACCTCTAGCA 4378
Qy 1759 GAACAACTACGTGAATGACTCAAAAGTATAGGCGAGGAGATACAAAGTTCAGCCATAC 1818
Db 4379 GAGCAGGTGAATAATTAACACATTAACGAGAGGCGCGGCTTACAGGTGACGGATAT 4438
Qy 1819 CACGTTAAGTCAATTTGACCAAGAAAGGACGCGTCCCTGTTCAGAACTTCCAGGCAAT 1878
Db 4439 GACGCGAGGCTCTACTACATGATGATCGGCATTCGGTCCCTGAGTTTCAAGCTTTG 4498
Qy 1879 AGTGAAGCGCTACATGCTTTTCAACGAGAGGAGTTGCTAAACAGATACCTGCACAC 1938
Db 4499 AGCGAGAGCCCACTATAGGTGTACAAAGAAAGGAGTTGTCTCAACAGAAACTATATCAT 4558
Qy 1939 ATCCATCAACCGAGGAGGCTTAAACATGACGAAAGATCTATAGACTGTAAAGACT 1998
Db 4559 ATTGCGCTCACGAGCCGTGCTGAACACGAGAGAACTACGAGAAAGTACAGACT 4618
Qy 1999 CAGGACACAGACTCAGAAATAGTCTTGCATTTGACCGACGAAAGTGTAAAGCGGAA 2058
Db 4619 GAAAGAACTGACCGCGAGTACGTCTTGCAGTATGATTAAGAAAGTCTGCTCAAGAGAG 4678
Qy 2059 GACGCAAGTCCCTTGTGCTTAAACCGGTATCTGGTATGATCCACTTTCAGAGTTTGG 2118
Db 4679 GAAAGCTGCGGTTGTGTGTGGAGAGCTAAACCAACCCCGTTCATGATTCGCC 4738
Qy 2119 TACGAGTCTCAAGACAGACCAAGCACTCAAGAAAGTCCCAACCATCGGAGCTAT 2178
Db 4739 TACGAGGCTGAAGATCAGCGCGTGGCACATTAAGACTACAGTATGAGGACTTT 4798
Qy 2179 GGAATGCGAGTTCAAGTTAATCTGAATCATCAAAAGCGCTGTGATTAAGAAAGATCTG 2238
Db 4799 GGGGTTCCGGATACGCAAGTCTGTATTTAAGGCTCTGTGACCAAAACAGATCTG 4858
Qy 2239 GTTGTAGTGCAGAAAGAAACTGCGCAGAAATCATCAGAGTGTAAAGAGATGAGA 2298
Db 4859 GTCAACAGCGGCAAGAGAGAACTGCCAGAAATATGTTAAACGACGTGAAGAACCGC 4918
Qy 2299 CGTATGATGTGTCTAGGACTGTGATTCAGTGTCTTAAATGGGGTTAAGCAACCC 2358
Db 4919 GGGAGGGGACAAAGTAGGAAACAGTACTCATCTGTCTAAACGGGTGTCTGTGGC 4978
Qy 2359 GTTAACTCTGTACATTTGATGAGGCAATTTGCTGCTGACATGAGGAGCGCTGTGGACTG 2418
Db 4979 GTGACATCTCTATATGTGAGAGAGCTTTCGCTTGCCATTCGCTTACTGTGCTGCGCTTA 5038
Qy 2419 ATTGCAATGCTCAACT--AAGAAAGGTATTTGTGCGGGAGCCCAAAACATGCGGC 2475
Db 5039 ATTGCTCTGTAAACCTCGAGCAAGTGTGTATGCGAGACCCCAAGCAATGCGGA 5098

2476 TTCTTAACTATGATGCTGTAAGTACATTTTAACTATGACATATGACATGAGTAC 2535
5099 TTCTTCAATATGATGCTGTAAGTACATTTTAACTATGACATATGACATGAGTAC 5158
2536 CATAAAGCATCTTACGAGGTGACACAGACTGTACCCGCTGTCTCCAGCTTCTC 2595
5159 CATAAAGTATATCCAGAGGTGACACAGACTGTACCCGCTGTCTCCAGCTTCTC 5218
2596 TACGACAGCGAATGAAGCGTTAACCACTGCTGATTAATATGATATGATACCA 2655
5219 TACGAGGCAAGATGCGACGACCAACCTGTCACCAACCAATATATCAACACCA 5278
2656 GGGACCAAGGCGGCAAGAGATGATCTTCACTGCTTTTACAGAGATGAGTAA 2715
5279 GAGACAGCAAGCGGCAAGAGATGATCTTCACTGCTTTTACAGAGATGAGTAA 5338
2716 CAGCTACGATTTGACATCAAAATACGAAATCATGACTGCGGTGATCGCAAGACTT 2775
5339 CAGCTGCACTGATGATGACGATGATGATGATGATGATGATGATGATGATGAT 5398
2776 AGCGGGAAGGCGTTTATGCTGCTGATGATGATGATGATGATGATGATGATGAT 2835
5399 ACCGCAAGGCGGATATGCGCTGATGATGATGATGATGATGATGATGATGATGAT 5458
2836 ACTTCTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2895
5459 GCGTGGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5518
2896 CTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2955
5519 CTGCGCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5578
2956 TCATGAGCACTGAGCGGCAAGAGCGGATGATGATGATGATGATGATGATGATGAT 3015
5579 ACATGAGCACTGAGCGGCAAGAGCGGATGATGATGATGATGATGATGATGATGAT 5638
3016 CAGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3075
5639 GCGCTGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5698
3076 GTCTTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3132
5699 GTCTTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5758
3133 TTCAAGCATGACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3192
5759 TTCAAGCATGACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 5818
3193 TTTGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3252
5819 TATGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5878
3253 GATGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3312
5879 AATGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5938
3313 AATGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3372
5939 GATGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5998
3373 GATGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3432
5999 GTTATGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6058
3433 GCGCGTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3492
6059 GCGCGTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6118
3493 GATGAGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3552
6119 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6178

3553 CCAGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3603
6179 TTGCTGAGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6238
3604 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3663
6239 TACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6298
3664 GTTGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3723
6299 ATTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6358
3724 AGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3783
6359 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6415
3784 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3843
6416 GTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6475
3844 AGTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3903
6476 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6535
3904 TTCTTGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3963
6536 TTCTTGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6595
3964 GACAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4023
6596 AGTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6655
4024 ATCAGAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4083
6656 AAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6715
4084 GGTCAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4143
6716 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6775
4144 AGACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4200
6776 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6835
4201 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4260
6836 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6895
4261 GCTGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4320
6896 GCTGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4380
4321 CCGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4400
6956 CCGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7015
4381 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4440
7016 CATGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7075
4441 CAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4500
7076 AGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
4501 GATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
7336 GATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7195
4561 AGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4620
7196 CGTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7255
4621 CATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4680

Accession	Sequence	Position
Db	7256 AACGAGCTGCTATTGTGATATGCGAGAGTACTGAGCTTGTGTGCCCACTGCGAAGGGCA	7315
Qy	4681 AATGAGCAGATTTGCTTGTATACATCCTGGGGAGAGTATGTCAGAGATCGCTCCAAATGC	4740
Db	7316 AACGAAACGATATGCTATACGCGCTGGGGCAAAACATGCAACATCATCATCCAAATGT	7375
Qy	4741 CCACTAGAGAGTCCAGAGCGCTCTGCTCACTCACACATTCATGCTGTGTAAATATAC	4800
Db	7376 CCGGTGAACGATTCGATTCATCAACACTTCCAGAGAACGAGCCCTGTGTGCGCTAC	7435
Qy	4801 GCTATGACGCGTGAAGCGGTATACAGTTGGCTCTGCGAAGAAAGAACAGTTGCGCGTA	4860
Db	7436 GCATGACGAGAAACGATCGCCCGCTTAGTGTACACCCAAAGTTAAAGATGTTGTT	7495
Qy	4861 TGTTCATCATTTCTGTTGCGGAGTACAGATCAAGGCGTGCAGAACTTCAAGTGCAGC	4920
Db	7496 TGTTCATCTTTTCCCTCCGAAATATCATGTAGATGGGGTGCAGAAAGTAAAGTGCAG	7555
Qy	4921 AAACCAAGCTCTTTTTCAGCGGTGTATCAACCGCTGTACACCCAGAAATTAAGCGGAA	4980
Db	7556 AAGGTTCTCTGTTGCAACCGACGATCTTTCAGTGGTTAGTTCGCGGAAATATGCGCA	7615

```

RESULT 9
US-09-190-246--4
; Sequence 4, Application US/09190246
; Publication No. US20030180257A1
; GENERAL INFORMATION:
; APPLICANT: Parrington, Mark
; APPLICANT: Li, Xiaomao
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
; FILE REFERENCE: Parrington et al.
; CURRENT APPLICATION NUMBER: US/09/190,246
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 8100
; TYPE: DNA
; ORGANISM: Semliki Forest virus
US-09-190-246-4

Query Match      29.5%; Score 1473.8; DB 12; Length 8100;
Best Local Similarity 57.9%; Pred. No. 0;
Matches 2779; Conservative 0; Mismatches 1982; Indels 36; Gaps 8;

QY      214  ATTGAAGTGGGCGCCGTGACATGACATTCACATCCACCGCTATCATTTGATCTGCGCT 273
Db       1  ATCGGCGATGGCGCCTTCCAGGAGAAATGATGTCATCGCAAAATACCACTGGTATGCCCT 60

QY      274  ATGATTAAGCGCTGAAGACCCGGACAGACTCAACAGGTATGCGAAGAACTTAAGAAA-- 331
Db       61  ATGGCGACCGGAGAAAGACCCGGAAGGCTCGATAGCTTACGAAAGAACTGGCAGCGGCC 120

QY      332  ---GTGACATTACCGACAAAGACATAGCCCTCTAAGGCGGAGACTGCTGGAAGTCATG 387
Db       121  TCCGGGAAGGTGCTGGAATAGAGATGTCAGAGAAAATACACGACCTGCAACCGTATG 180

QY      388  TCAACACCCAGACGCGAGAGACTTCATCTTGTGTATGACACAGACGCCAGGTGAGGTAC 447
Db       181  GCTACGCGCAGACGCTGTAATCTTCTACCTTTTGCCGTGACATAGACGTCAGTTCGTA 240

QY      448  TTGGAAGTAGACAGTATACCAAGATGTAAGAGAGTCATGACCGGACATCAATCTAC 507
Db       241  GCAAGCCGAGTGGCCGTATACAGGACGCTGATGCTGACATGACCAACATCGCTGTAC 300

QY      508  CACCAAGCGCTTAAGAAGAGATTAGACAATTACTGATAGAGCTTTGACACGACCCCTTT 567
Db       301  CATCAGGCGATGAAGAAGTGTCAAGAACGCGATTGATTTGGGTTTGACACACCCCGTTT 360

QY      568  ATGTACAAAAACATGCGAGTTTCTTACCTTACCTTAACAACAGAACTGGGCTGACGAGAG 627

```

Db	361	ATGTTTGACGGCTAGCAGCGCCCGTATCAACCTACGCCACAAACTGGGCGGACGACGACG	420
Qy	628	GTAATGGAGCAGCTAACATTGGCCTCGGTAACTCAGATCTTCAGAGAGCAGGCTTGG	687
Db	421	GTGTACAGCGCCAGGACATAGAGACTGTGTGACAGACTCTTGACTGAGGAGAGACTCGGC	480
Qy	688	AAACTCTCAATCCTTAGAGAGAGAGGCTCCAACTACTAATTAAGTATATTTCCGGT	747
Db	481	AAACTGTCAATTCCTCCGMAAGCAATTGAAACCTTCGCACACGTATGTTCCGGTA	540
Qy	748	GGTTCAACAATCTACACAGAGATAGATCACTGTTACCTAGCTGGCATCTTCCAAACGTG	807
Db	541	GGATCTACATTTGTACACTGAGACGAGAAAGCTACTGAGAGCTGGCACTTACCTCCGTA	600
Qy	808	TTCCACTTGAAGGAAAGCTTAACCTTCACAGGTAGATGTGGAGCCATTGTACGTTGAA	867
Db	601	TTCCACCCTGAAGGTAAACAATCCTTACCTGTAGTGCAGTATCATGTGTAATCATGTGAA	660
Qy	868	GGGTACGTCATCAAAAAAGATTAACGATCGCCACGACATATACGTAAGATTGAGAAACTTG	927
Db	661	GGGTACGTAGTTAAGAAATATCAATATGTGCCCCGCTGTACGTTAAGTAACGGTAGAGTAC	720

Db	721	GCCTGACGTATCACGGGGAGGATTCCTAGTGTCAAGACCAACGACACTGTCAAAGGA	780
Oy	988	GAGAGGGTTTCCTTTGCTGTGTGTACGTATGATACAGCCACACTTTGGCATCAGATGACA	104
Db	781	GAAGAGCTCATTCCTCCGTATGACCACTACCTCCCTCAACCATCTGTGATCAATAGACT	840
Oy	1048	GGGATTCTGGCACTGAGCTTATGTGTGATATACCGACAAAACCTTTGGTTGGGCTCAAC	110
Db	841	GGCATCTACGAGCACCGACGTACAACCGGAGACGACAGAAGTTGTGTGTGGATTTGAAT	900
Oy	1108	CAAGGATTTGTGTCAATGTATGGACGCGAAAGAAATATTAACACATGACAGAACTATCTA	116
Db	901	CAGAGGATAGTTGTGTGAACGGAAGAACACAGCGAAACATTACCATGTAAGAACTATCTG	960
Oy	1168	TTACAGTGTGTGCGCCAGGCGTTTTTCCAGGTGGGGCGGTGAACATCGTCCGACTTGGAC	122
Db	961	CTTCCGATTTGTGGCCGTGCGCATTTTAGCAAGTGGGGAGGAATACAGGACAGACTTGAT	102
Oy	1228	GACGAGAAAGAACTAGGGGTGCGGGAGCGCACTCTTACTATAGGTGCTGCTGGGCTTTC	128
Db	1021	GATATAAAACCTCTGGGTGTCCGAGAGAGTCACTTACTTCTGCTGTGTGGGCAATTT	1080
Oy	1288	AAGACCCAGAAATACATCTCATCTACAAAGAGCTGTAGCGAAACATTAAAGAAAGTA	134
Db	1081	AAAAAGAGAAAGATGACACACCATGTACAGAAACAGACACCCAGACAAATAGTAGAGTG	114
Oy	1348	CCTGCGCTCTTTGACTATTTGTGATTTCCAGGCTTTACAGCCACGCGCTCGATATGGGC	140
Db	1141	CCTTCAGAGTTTAACTGTGTGTATCCGAGCCATAGTGTACAGGCTGTGCAATCCCA	1200
Oy	1408	TTCCGCGGTAGCTCAAGCTGCTCTCTTGAACCAACTGTCAAAACCCGACCGGCTATTTACA	146
Db	1201	GTCAGATCACCCATTTAAGATGCTTTTGGCCAAAGAACCAAGCGAGA--GTTAATACCT	125
Oy	1468	ATGCCCATGTGAGCATCTGTGGGTGCTTAAAGCAAGAAAGCTMAAGATGTGCTGCACAGG	152
Db	1258	GTTCTCGACGGGTGTCAAGCCAGGAGATGCTAAACAAGGAGAAAGAGATTTGAGGCG	131
Oy	1528	GAAGAGATTCAGAGAAGCCCTGCCACCTTTGCTCCCTGAATATGAAAAAGAGACCGTAGAG	158
Db	1318	GAGCTGACTAGAGAAAGCTTACACACCTCGTCCCATTCGGCGCGGCGGAGACGGGAGTGC	137
Oy	1588	GCAGAAGTAGACCTCATTTATGCAAAG-----GCAGAGCAGAGTAGCGTGTGAGACACCA	164
Db	1378	GTCAGCTGTGAGCTTTGAAGAACTAGTAGATACACGAGGTGACGGGGGTGTGTGAAACACCT	143
Oy	1642	CGAGGACACATCAGGGGTGCAAGTTTCCCAAGCCGAAGAAAGATTGGGTCTTACGCTATA	170

Db 1438 GCAGCGCGTTGAAATGACGCGACGCGAAACGATCTAGTAAATTAACGTAGT 1497
Qy 1702 CTTTCAACCCAGGCGGTATTTGAAATAGTGAATAAATGCGGTGATTCACCCATTTGGCGGA 1761
Db 1498 CTGTCCCGGACGACCGTCTCTCAAGAGCTCAAGTTGGCCCCGTCACCTCTTACAGAG 1557
Qy 1762 CAAGTACTGTGAATGACTCACAAAGGTAGGGGAGGAGATACAAAGTCAGGCTATACAC 1821
Db 1558 CAGGTAAATATATTAACATTAACGAGGAGGCGCGGTTACCAAGTCGACGATATAGAC 1617
Qy 1822 GGTAAAGTATTTGATCCAGAAAGGAGCGCGCTCTGTTCAAGACTTCAGGCAATTTAGT 1881
Db 1618 GCGAGGGTCTTAATCAATGATGATCGGCAATTCGGTCCCTGAGTTTCAAGCTTTGAGC 1677
Qy 1882 GAGAGCGCTAGCATCGTTTTCAGAGAGGAGTTCGTAACAGATACCTGACACCATC 1941
Db 1678 GAGAGCGCTAGCATTTGATGATCAACGAAAGGAGTTTCGTAACAGGAACTTATCATATTT 1737
Qy 1942 GCAATCAACGAGGAGCGCTAAACATGACGAGAGTACTATTAAGACTGTAAAGACTGAG 2001
Db 1738 GCGGTTCACGAGCGGTGCTGACACGAGCGAGGAGTCTAGAGAAAGTCAAGACTGAA 1797
Qy 2002 GACACAGACTAGAAATACGTCTTGATATTTGACGACGAAAGTGTGTTAAGCGAGAGAC 2061
Db 1798 AGAATGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1857
Qy 2062 GCAGGTCCCTGTGCTGATACGCGTGAATGATGATGATGATGATGATGATGATGATGATGAT 2121
Db 1858 GCGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1917
Qy 2122 GAGAGTCTCAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 2181
Db 1918 GAAAGGCTGAAGATCAAGGCGCTCGGACCATATTAAGATCAAGATGATGATGATGATGAT 1977
Qy 2182 GTGCCAGTTCAGTAAATCTGGAATCATCAAAAGCGCTGTGATGATGATGATGATGATGAT 2241
Db 1978 GTTCCGGATCAGGACAGTCTGATTTAAGAGCTCGTGACCAACGATCTGATC 2037
Qy 2242 GTGAGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2301
Db 2038 ACCAGCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2097
Qy 2302 ATGATGTTGCTGCTAGGACTGTGATTCAGTGTCTTCAAAATGGGTTAAGCACCCTTC 2361
Db 2098 AAGGGGCAAGTATGAGGAAACAGATGATCTCATCTGCTAAACGGGTGCTGTGCTGCTG 2157
Qy 2362 AACACTGTATCATTTGATGAGGCAATTTGCTGCAATGACGAGGACGCTGCTGCACTGAT 2421
Db 2158 GACATCTATATGATGAGAGAGGCTTTCGCTTGCATTCGGGATCTCTGCGCTTAATTT 2217
Qy 2422 GCCATCGTCAAACT--AAGAAAGTGTATTTGCGGGAGACCCAAACATGCGGCTTC 2478
Db 2218 GCTCTTTGTAACCTCGGAGCAAGTGTATGTCGAGAACCCCAACCAATGCGGATTC 2277
Qy 2479 TTTAATCATGATGCTGCTGAAGTACATTTTAACCTGATATGATGATGATGATGATGATGAT 2538
Db 2278 TTTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2337
Qy 2539 AAAAGCATCTTGAAGAGTGAACAGATGATGATGATGATGATGATGATGATGATGATGAT 2598
Db 2338 AAAAGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397
Qy 2599 GACAGGCAATGAGAGAGGTTAATCCATGCTGATTAATATCATATGATGATGATGATGATGAT 2658
Db 2398 GAGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2457
Qy 2659 ACCAAGAGCGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2718
Db 2458 CAGAGCAAGCGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2517
Qy 2719 CTACAGATTTGATCAAAATCAAGAAATCAATGATGATGATGATGATGATGATGATGATGAT 2778
Db 2518 CTGAGATTTGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2577

Qy 2779 CGAAGAGCGTTTATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2838
Db 2578 CGCAAGAGGATATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2637
Qy 2839 TCTGAGCAGTGAACGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2898
Db 2638 TCGAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2697
Qy 2899 GCTGTGATCCCTGATTAAGACATTTACCTTATGATGATGATGATGATGATGATGATGATGAT 2958
Db 2698 GCGGCGATCCCTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2757
Qy 2959 TTTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3018
Db 2758 TTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2817
Qy 3019 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3078
Db 2818 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2877
Qy 3079 TTTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3135
Db 2878 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2937
Qy 3136 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3195
Db 2938 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2997
Qy 3196 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3255
Db 2998 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3057
Qy 3256 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3315
Db 3058 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3117
Qy 3316 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3375
Db 3118 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3177
Qy 3376 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3435
Db 3178 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3237
Qy 3436 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3495
Db 3238 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3297
Qy 3496 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3555
Db 3298 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3357
Qy 3556 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3616
Db 3358 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3417
Qy 3617 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3676
Db 3418 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3477
Qy 3677 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3736
Db 3478 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3537
Qy 3737 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3796
Db 3538 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3594
Qy 3797 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3854
Db 3595 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3654

QY	3847	TTTACCGGTGCTGTCAAGCCCTAAGAACACTGCGGAAATACTGAGGTTCTCTGTGTC	3906
Db	3655	TTGCAAGAGTGTGCGCCGGGATTGTGTCAACAGCAATTAAGAAGTGTCTTCTGTTCC	3714
QY	3907	TTGCGCAAGGCAACGGCAACCAACACATGATACAGAGACAGACTCGGTGTAGTCTTGAC	3966
Db	3715	TTCAAATTTTGCAACGGAAAGAGNACCTCTACGCTACCAAGATGAATACAACTAGT	3774
QY	3967	AACATCTATCAAGGCTCAACGAGTACGAGGACGGAGAGCTCAACGCTACAGAGTATC	4026
Db	3775	GCCGTGTATGCGGAGAAAGCCATCAACGCGCGGGTGTGCAACATCTTAAGAGTTAAG	3834
QY	4027	AGAGGTGACATTACCAAGAGCGGCTGACCAACTATTCGTTATGTGCTAATAGCAAGGT	4086
Db	3835	AGACGACACATAGCCACGTGACAGAAAGCGGCTGTGTTAACGACGCTTAAGCCGTGGA	3894
QY	4087	CAACGAGTTTCGGAGTGTGCGGTGCACTGTACCGAAATAGGCGGCTGCTTTGATAGA	4146
Db	3895	ACTGTAGGGGATGGCGTATGCAAGGCGCTGGCCGAAGAAATGGCGCTTAAGGA	3954
QY	4147	CAGCCAATAGCTGTGCGGACGGCTAGACTTGTGAAGACGAACCG--CTCATATACAT	4203
Db	3955	GCAGCAACACAGATGGGCACAATTAAACATGTCGCGCTGTACCCGCTATCCAC	4014
QY	4204	GCTGTAGACCCCATTTTTTCTTAATATGCGGAAACCGAGGGCGACCTTAACCTGCACT	4263
Db	4015	GCTGTAGGCGCTTAATTTCTGTGCCACGACTAAGGGAAAGGGAGACCGGAATTTGGCGCT	4074
QY	4264	GCCATACATGACATAGCGTCCATCGTCAACGCTGACGGATTAACAAAATATCAGTACCG	4323
Db	4075	GTCATCCGGGAGATGGCCCGGAAAGTAAACAAGCTGTCACTGACAGGTAGCCATCCCG	4134
QY	4324	CTACTGTCAACCGGCATCTATTCTGTGTGGCAAGATGAGTATGCAATCAATTGCATCAC	4383
Db	4135	CTGCTGTGCCAGAGTGTTCAGCGGCGGAAAGATAGCTGACAGCAATCCCTCAACCAT	4194
QY	4384	CTGTTTACCTGTTTCGACATACGATGCGGATGACCAATATATGTTGGATTAACAA	4443
Db	4195	CTATTTCACGAAATGAGCCACGACGCTGACGATCTACTGTGAGAGACAAAAGT	4254
QY	4444	TGGGAGACCAAGATTAATCGAGGCACTTACCCGAAAGAAACGTGCAAAATCTGGATGAT	4503
Db	4255	TGGGAGAAAGAAATTCAGGAAGCCATTGACATGAGGACGGCTGTGAGATTGCTCAATGAT	4314
QY	4504	GACAAGCCAGTAGACATTGACTTGGTCAAGGGTCCACCCAAACAGCTCTTTGGACGACGA	4563
Db	4315	GACGTGAGCTGACACACAGACTTGGTGAAGTGCACCCGGAACGAGCTGGTGGTGTGT	4374
QY	4564	CCAGGTTACTCCGTCATATGAGGGCAAGTTGTAATTCATACCTGGAAGGTACACGATTCAT	4623
Db	4375	AAGGGCTACAGTACCACTGACGAGGGTGCCTGTATCTGTACTTTGAAGGTACAAAATTCAC	4434
QY	4624	CAGACCGGCAAGGACATGCGGAAATCCATGCAATGTGGCCCAACAAATCTGAGGCTAAT	4683
Db	4435	CAGGCTGTATTGATATGACAGAGATTACTGACGTTGTGTGGCCGACACTGCAAGAGGCAAC	4494
QY	4684	GAGCAGATTTTCTGTACATCTGTGGGAGAGATATGTCACAGCATCCGCTCCAAATGCCA	4743
Db	4495	GAAACGATATGCTATATACGCGCTGGGCGAAACAATGACAAATCATCATCAATCAATGTCG	4554
QY	4744	GTAAGAGAGTCAGAGGCGTCTGTCTCACTCAACACTTCACTGCTGTGTATTAAGCT	4803
Db	4555	GTTGAACGATTCGGATTCTATCAACAACCTCCAGAGACAGTGCCTGTCTGTGCGCTACCA	4614
QY	4804	ATGACCGGTGACCGGTATACAGGTTGGGCTGTGGAAGAAAGAACAGTTGGCGGTATGC	4864
Db	4615	ATGACAGCAGAAACGATGCCGCTTGAAGTACCAAGTTAAAGATGGTGGTTTGC	4674
QY	4864	TCATCATTTCTGTTGCGCAAGTACAGAGTACACAGCGCTGACGAAGCTACAGTCAGCAAA	4923
Db	4675	TCATCTTTTCCCTCCGAAATACATGTAAATGAGGGTGTGCAAGAGTAAAGTGCAGAAAG	4734
QY	4924	CGAGTCTGTTTTAGGCGCTGTACACCGGCTGTACACCCGAGAGTACCGGAA	4980

DB	4735	GTCTCTCTGTTGCAACCCGACGAGTACTTGTAGTGGTTAGTTCGGCGGAGAGATGCGCA	4791
	RESULT 10		
	US-09-507-362-102		
	Sequence 102, Application US/09507362		
	Publication No. US20030096397A1		
	GENERAL INFORMATION:		
	APPLICANT: Dubensky Jr., Thomas W.		
	Polo, John M.		
	Belli, Barbara A.		
	Schlesinger, Sondra		
	Dryga, Sergey A.		
	Frolov, Ilya		
	TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS		
	WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR		
	SYNTHESIS		
	NUMBER OF SEQUENCES: 125		
	CORRESPONDENCE ADDRESS:		
	ADDRESSEE: Seed Intellectual Property Law Group PLLC		
	STREET: 701 Fifth Avenue, Suite 6300		
	CITY: Seattle		
	STATE: Washington		
	COUNTRY: USA		
	ZIP: 98104-7092		
	COMPUTER READABLE FORM:		
	MEDIUM TYPE: Floppy disk		
	COMPUTER: IBM PC compatible		
	OPERATING SYSTEM: PC-DOS/MS-DOS		
	SOFTWARE: PatentIn Release #1.0, Version #1.30		
	CURRENT APPLICATION DATA:		
	APPLICATION NUMBER: US/09/507,362		
	FILING DATE: 18-Feb-2000		
	CLASSIFICATION: <Unknown>		
	ATTORNEY/AGENT INFORMATION:		
	NAME: McMaisters, David D.		
	REGISTRATION NUMBER: 33,963		
	REFERENCE/DOCKET NUMBER: 930049.457D6 /1196.011		
	TELECOMMUNICATION INFORMATION:		
	TELEPHONE: (206) 622-4900		
	TELEFAX: (206) 682-6031		
	INFORMATION FOR SEQ ID NO: 102:		
	SEQUENCE CHARACTERISTICS:		
	LENGTH: 8000 base pairs		
	TYPE: nucleic acid		
	STRANDEDNESS: single		
	TOPOLOGY: linear		
	SEQUENCE DESCRIPTION: SEQ ID NO: 102:		
	US-09-507-362-102		
	Query Match	26.0%; Score 1300.4; DB 11; Length 8000;	
	Best Local Similarity	56.0%; Pred. No. 0;	
	Matches 2819; Conservative	0; Mismatches 2131; Indels 80; Gaps 16;	
QY	23	ATATGAAAGAAATTCACGTTGACTTTCAGTGTGACGACGCCGATATGTCAGAGTGTATACAC	82
DB	64	AGAAGCCAGTAGTAAACGTAGACGTGAGACCCCAAGGTCCGTTTGTCTGCAACTGCACAA	123
QY	83	GGAAGCTTCCAAATTTGAGATGGAAGCAAGCAGGATCACTGACATGACATGACATGCAATG	142
DB	124	AAAGCTTCCGCAATTTGAGGTAGTGTAGCAGACAGAGGTCACTCCAAATGACCAATGCTAATG	183
QY	143	CCAGAGCGTTTTCGATGTGTGCAACAAAGCTCATTTAGAGCGAAGTGCACCGGACCAAG	202
DB	184	CCAGAGCATTTTTCGATCTGTGCACTGCAAGTAATGACAGCTGAGAGTCTTCTTCAACACAGCA	243
QY	203	TTATCTTGAACATTGGAAGTGCAGCCCGTCAGACATGACATTCATCAATCAACCGCTATCAT	262
DB	244	CGATCTTGACATAGGACAGCGCACCGGCTCGTAGAATGTTTTCGAGCACCAAGTATCAT	303
QY	263	GTATCTCCCTATGATAGCGCTGAAGACCCGGAACGACTACACCGTATGCAAGAAAGC	322

Db 304 GTGTCTGCCCAATGCGTAGTCCAGAAAGCCCGACCGCATGTGAATATGCCAGTAATAC 363
Qy 323 TTAAGAAA-----GTGACATTACCGACAGAAACATAGCTCTTAAGGCGGAGACTGCG 376
Db 364 TGGCGGAAAAGCCGTGCAAGATTACAAAGAAAGAACTTGATAGAAAGATTAAAGATCTCC 423
Qy 377 TGGAGTCATGTCACACCGACGAGAGACTTCATCTGTGTATGACACAGACGCCA 436
Db 424 GGACGGTACTTATGCGCCGATGCTGAAGACATCGCTGTGCTTTCAACAGATGTTA 483
Qy 437 CGTGTAGTACTTTGGAAGTGTAGACATACCAAGATGTGTACGACGTCCATGACCGA 496
Db 484 CCTGCAACATGCGTCCGATATATCCGTATGACGAGACGTGTA---TATCAACGCTCCCG 540
Qy 497 CATCAATCTAACACGAGCGCTTAAAGAGTAAAGCAATTTACTGATAGGCTTTGCA 556
Db 541 GAACTATCTATCATAGGCTATGAAAGGGTGGACCTGTATGATGATGGCTTCGACA 600
Qy 557 CGACCCCTTTTATGTACAAAACATGCGAGTTCCTACTACTTAAACAACGAACTGGG 616
Db 601 CCACCCAGTTCATGTTCTCGGCTATGCGAGTTCGTACCTGCGTACAAACCAACTGGG 660
Qy 617 CTGACGAGAGATTTGGAAGCATTAATGCGCTCGGTATCTAGATCTTTCAGAGA 676
Db 661 CCGAGAGAAAGTCCCTGGAAGCGGTAAACATCGACTTTGACAGCAAAAGCTGAGTGAAG 720
Qy 677 GCAGCTTGGAATCTCAATCTTAGAAGAAAGAGCTCCAACTACTAATTAAGTCA 736
Db 721 GTAGGACAGAAATTTGTGATTAAGAGAAAGAGTTGAAGCCGGGTCGGGTTT 780
Qy 737 TATTCCTGCTTGGTTCACCAATCTACACAGAAATAGATCACTGTTACGTAGCTGGAC 796
Db 781 ATTTCTCGTGAGATGACACTTATCCAGAACAGAGCCAGCTTGACAGACTGGACATC 840
Qy 797 TTCAAAGCTGTTCACCTTGAAAGAAAGTCTTAACCTTACAGATGATGTGGACATTG 856
Db 841 TTCCATCGGTGTCCACTTGAATGGAAGCACTGCTACACTTGGCTGTGATACATGCG 900
Qy 857 TCAGCTGGAAGGGATACGTCATCAAAAGATTAACATGAGCCGACGATACATACGTTAAG 916
Db 901 TGAGTTGGAGAGGCTACGTAAGTGAAGAAATACACATGCTCCGGAGTACACGGAGAAA 960
Qy 917 TTGAGAACTTGGCGTCCACATGCAATCGGAGGGTTCCTTGAGTTGAAAGTCAAGATA 976
Db 961 CCGTGGATACCGCGTTACACACATAGGAGGCTTCTTGCTATGCAAAAGTTACTGACA 1020
Qy 977 CGCTGCGCGGAGAGGGTTCCTTGTGTGTGTAGCTATGATACGACGACACTTGGG 1036
Db 1021 CAGTAAAGAGAAAGGGTATCGTTCCTGTGTGACGTAATCCCGGACCATATGCG 1080
Qy 1037 ATCAATGACAGAGATTTCTGCACTGACGTTAGTGTGATGACGCAAAAACATAATGG 1096
Db 1081 ATCAATGACAGGATTTATGAGGACGAGATATCACTGACGATCCAAAACCTTCTGG 1140
Qy 1097 TTGGGCTCAACCAAGATTTGTCTCAATGTAGAGCGCAAGAAATATAACAAATGC 1156
Db 1141 TTGGGCTCAACCGAGATTTGATTAACGTAAGACATAACAGAAACCAACACATATC 1200
Qy 1157 AGAATCTATCTATTCAGTGTGCGCCAGCGCTTTTCCAGGTGGGCGCGTGAATCGTG 1216
Db 1201 AAAATTACTCTTCTGCGATCATAGCACAAGGGTTCAGCAATGGGCTAAAGAGCCCAAG 1260
Qy 1217 CCGACTTGGACGACGAGAAAGAACTTAGGGGTGCGGAGCGCACTCTTACTATGGGCTGT 1276
Db 1261 ATGATCTTGATTAACAGAAATGCTGGGTACTAGAGAAAGCAAGCTTACATGATGCTGT 1320
Qy 1277 GCTGGCTTTCAAGACCCAGAAATCAACATCATCTTCAAGAGCTGGTATGCAAAACA 1336
Db 1321 TGTTGGCGTTTGCACTAAGAAAGTACATCTGTTTATGCGCCCACTGGAACGACAGCT 1380
Qy 1337 TTAAGAAAGTACCTGCGCTCTTGTGACTATTTGTGATTCACAGCCTTACAGCCACGGG 1396
Db 1381 GGTAAAGTCCCAAGCTCTTTTAAAGGCTTTTCCATGTGTCGATGAGACGACTCTT 1440

Qy 1397 TCGATATGAGGCTTCCGCGTAGAGCTCAAGCTGCTGTTGAACCA---ACTGTCAAACCG 1453
Db 1441 TGCCATGTCCGTGAGGACAGAAATTTGAATCTGACATTGCAACCAAGAAAGAGAAAAAC 1500
Qy 1454 CACCGCTATTAACAATGCGCGATGTGAGCATCTGCGTGGCTTACAGCAAGAGCTGAAG 1513
Db 1501 TGCTCAGAGTCTCGGAGGAATTAATGTCATGAGGCGCAAGGCTGCTTTTGAAGATGCTCAG 1560
Qy 1514 AAGTGTGACGCGGAAAGATCAGAGAAAGCCCTGCAACCTTGTCTCCCTGAATATGAAA 1573
Db 1561 AGGAAGCCAGAGCGGAGAAAGCTCCGAGAAAGCACTTCCACATTAATGTGCGACGAAAGCA 1620
Qy 1574 AAGAGACGTAGAGCA-----GAAATGACCTCATTAATGCAAGGCGAGGAGCAG 1624
Db 1621 TCGAGGACCGCGACAGATTTGCTCGCAAGTGTGAGAGGGCTCCAGCGGACATCGAGACAG 1680
Qy 1625 GTAGCGTGAAGACACACAGAGACACATCAGGCTGACAAATTAACAGCGAAGAGAGA 1684
Db 1681 CATTAATTTGAACCCCGGCGGTGACGTAAAGATTAATCTTCAAGCAATGACCGTATGA 1740
Qy 1685 TTGGCTTTTACGCTATCTTTCACCCAGCGCGTATTTGAATGTGAATAAACTGGCTGTA 1744
Db 1741 TCGGACATATATCTGTGCTCGCCAACTGTGTGTAAGAAATCCAAACTCGCACAG 1800
Qy 1745 TCCACCAATTGGCGGAAACAAGTACTGTAAATGACTCAAAAGGTAGGCGAGAGATCA 1804
Db 1801 CGCACCCCTGACCAATCAGTTAAGATCATTAACATCTCGGAAGATCGAAGAGTACG 1860
Qy 1805 AAGTGAACCATATCACCGGTAAAGTCAATTTGTAACGAAGAGACGCGCTCTGTTCAAG 1864
Db 1861 CGGTGAACCATATGAGCGCTAAAGTATGATGCGACAGAGAGTGCCTGATCATGCGCAG 1920
Qy 1865 ACTTCAAGCATTTGAGTGAAGCGCTATGATGCTTTTCAAGAGAGGAGTTTGTAAACA 1924
Db 1921 AATTCCTGACACTGATGTAAGCGCACGCTTAATGTAACAGAAAGAGTTTGTGAACC 1980
Qy 1925 GATATCTGACACATCTGCAATCAACGAGAGGCGCTAAACATGACGAAAGTACTTA 1984
Db 1981 GCAACCTATACCATCTGCTCATGATGAGCCCGCCCAAGAAATACAGAAAGAGAGTACA 2040
Qy 1985 AGACTGTAAAGACTCAGAGACAGACTGAGATAAGTCTTGTGATTAATGACGACAAAGT 2044
Db 2041 AGGTTACAAAGGCAAGCTTGACAGAAACAGATGATGTTGACGTGGAACAAAGCGTT 2100
Qy 2045 GTGTTAAGCGAAGAACCCAGGTCCTTGTGTCTTAACCGGTGATTTGTGATCAACAT 2104
Db 2101 GCGTTAAGAAAGAAAGACCTCAGGTCGTGCTCTCGGAGAACTGACCAACCTCCCT 2160
Qy 2105 TTCAAGAGTTTTCGTAAGAGTCTCAGAGACGACGACGACGACCTCACAAGTCCAA 2164
Db 2161 ATCATGACCTAGCTCTGAGAGACTGAAGACCCGACCTGCGGTCCGTAAGAGTCAAAA 2220
Qy 2165 CCATCGAGTCTATGAGTGCAGGTTCAAGTAACTGAAATCATCAAAAGCGCTGTA 2224
Db 2221 CAATAGAGATGATAGGACACCGGGGTGCGGGAATCAGCTATTAATCAAGTAACTGTCA 2280
Qy 2225 CTAAAGAAATCTGTTGTGAGTGCAGAAAGAAAGAACTGCGCAGAAATCATCAGGATG 2284
Db 2281 CCGCAGAGATCTTGTATCCAGCGGAAAGAAAGAAATTTGCGCAAAATTTGAGGCGGACG 2340
Qy 2285 TAAAGAGATGAGAGCTATGATGTGTGCTGTAGACCTGTGATTCAGTCTTCAATG 2344
Db 2341 TGTAAAGACTAGGGGATGAGATTAAGTGAACAGATGATGATGCTTAATGCTCAACG 2400
Qy 2345 GGGTTAAGCACCCGTTTAAACTCTGTATCAATTAATGAGGCAATTTGCTGCAATGAGGGA 2404
Db 2401 GATGCCAAGAGCGGTAAAGTGTGATGCTTGAAGAAAGCGTTCGCGTGAACGAGAG 2460
Qy 2405 CGCTGCGCACTGATTTCCATGCTCAAC---TAAAGAAAGTGTATTTGCGGGAAC 2461
Db 2461 CACTACTTGCTGATGCTATGCTGACAGCCCGCAAGAAAGTATGATATGTCGAGAC 2520

OY	2482	CAAAACATGCGGCTCTTTTAA	CATGATGTGCTGAAGTACATTTTAA	CC-----	2512		
Db	2521	CCATGCAATGGGGATTCTTCAA	CAAGAATGCAACTAAAGTCACTTTCA	ACCCCTGAA	2580		
OY	2513	ATGACATATGCACTGAAGTGA	CCCTAAAGATCTCTAGAGGGTGA	CACAGACTGTA	2572		
OY	2573	CCGCGCATGCTCTCAACGCT	CTTCTTACGA	CAAGCGAATGAAGACGGTTA	ACCATGTGCTG	2632	
Db	2641	CAGCTATATGTATCGACA	CTGCACTTACGATGGAAAGATG	AAACCA	CGAACG	2700	
OY	2653	ATTAATATCATCATAGATAT	CCACAGGACCA	CAAAAGCCGCA	AAAGATATCTGATTTCTTA	2692	
Db	2701	AGAACTATGAATTCATATAT	ACAGGGGCA	CAAAAGCCGCA	AGCGGATATCATCTCGA	2760	
OY	2693	CCTGTTTCAAGAGATGGGTGA	AAACGCTAC	GATTGAC	CTAACAAAATACGAAATCATGA	2752	
Db	2761	CATGTTCCGGGGTGGGTTA	AGCAATTGCAATTC	CCGACATGA	GTAAATGA	2820	
OY	2753	CTGCGGCTGCACTCGCAAG	ACTTACGCGGAAAGCGTTAT	GTCTGCTCAGTACAAAGTCA	2812		
Db	2821	CAGCGCGGCGCTCA	CAAGGGCTTAC	CAAAAAGAGTGTAT	GTGCGTCCGGCAAAAAGTCA	2880	
OY	2813	ACGAGATCCCACTCTAC	CTCCAGACTTTCTG	AGCACTGAACGTGTACTTAC	AGCCACAG	2872	
Db	2881	ATGAAAACCCACTGTAC	CGCATCATCAGAC	CATGTAA	CGTGTGCTCACCCGCACTG	2940	
OY	2873	AAAAACGACTGTCTGGAAG	ACCGCTAGCGGGTGA	TCCGTGGTAAAGCA	CTTACAGCTA	2932	
Db	2941	AGGACAGGCTAGTGTGA	AAACCTTGCAGGGGCA	CCATGGATTBAGAGCCCACTTACA	3000		
OY	2933	AATATCCCGGGATTTCA	CGGCTTCATTGAG	CGACTGACGCGGCA	CACGACGCATTA	2992	
Db	3001	TACCTAAAGGAAAC	CTTTCAAGGTACTAT	TAAGAGACTGGGAAGCTGA	CACAAAGGAATTA	3060	
OY	2993	TGGCAGCGGTTCTTAT	TAAGCCGACAGAC	GTGATGTTC	CCAGATTAAGTGAACGTCT	3052	
Db	3061	TTGCTGCATATAAC	CAAGCCCACTCCCGTCC	CAATCCCTTACGTGCA	CAAGCAACGTTT	3120	
OY	3053	GCTGGGCGAAGGCTTTAG	AGCCAGCTTGGCCACGGGCA	CACTTGTGCTGA	CGAGACAGC	3112	
Db	3121	GCTGGGCGAAGCATTTGA	ACGATTA	CTAGCCAGGGCGGTAT	GTCTTACCGTTGCGC	3180	
OY	3113	AGTG- - -GAGAC	GTTCGACCCATTTCA	AGCATGACAGAGCGT	ACTACCTGAATTGGCAC	3169	
Db	3181	AGTGAAGGA	ACTGTTCCCA	CAGTTTGGGATG	GAACAACAATTCGGCCATTTACGCT	3240	
OY	3170	TGAATCTCTTTTGC	CAACAGGTTCTTTGA	GTAGAC	CTGGAACGTGGTATTTTCCGCTC	3229	
Db	3241	TAGACGTATTTTGC	ATAAGTTTTTGG	CACTGAC	TTGAACAAGCGGCACTGTTTTCTAAAC	3300	
OY	3230	CTACCGTGC	CTTACTTACAG- - -	-----	GGATAGACATGGGATA	3268	
Db	3301	AGAGATCCCACTAA	CGTACCATCCCGGAT	TACGAGGCGCGT	AGCTCATTTGGGACA	3360	
OY	3269	ACTCGCCAGGGA	GAACA	CACTGATGGGCTTTAT	ATAGAGAGT	AGCAAGAGTGTCA	3328
Db	3361	ACAGCCAGGA	AAACCGGCA	AGTATGGGTAC	GTACAGCACTTCCGCGCA	ACTTCCGTA	3420
OY	3329	GATATCGGTGC	ATCAAAAGCGGTTGA	CACAGGCA	GGGTAGCTGATTAAGAAATATA	3388	
Db	3421	GATTTCCGGTGTTC	AGCTAGC- - -	TGGGAAGGGGAC	ACA	CACTTTTGA	3477
OY	3389	CCATTAAGAG	CTACTCTCC	ACAATTAATGTGTTC	CA	TTAAATGCGCGTTGCCCACT	3448
Db	3478	GAAACGAG	GTATCTCTG	CACAGCATTA	CTGTGTCCGGTGA	ACGCAATCTTCTCA	3537
OY	3449	CGTTATG	GTGACACAAAGGAC	AGGGAACA	ATGATCA	CAGCGGATTTCCATCTTA	3508
Db	3528	CTTATGT	CCCCGAGTAC	AAAGAGCA	AAACCGCGCGGTCA	AAAAAATTTCTTGA	3597
OY	3509	TGAAGGCA	AAATCTGTGTGTG	ATGATC- - -	GGGATCTTAT	CAGCATTCAGGGAAGAAAG	3565

Db	3558	TCAAACACCACTCAGTACTTGTGGATTCAGAGAAAAATTGAAGCTCCCGTAAAGAA	3657
Oy	3566	TAGAGTCATGGGTCCATTGGCCCA-----TAATACCATCAGGTGTGATCTCGATTGG	3619
Db	3658	TCGATGATGATGCCCCCAATTGGCATTACCGCGGTGCAATPAGAATTACAACTGGCTTTGC	3717
Oy	3620	GAATACCTAGCCATGTCCGTAAATATGACATTATCTTTGTCATGTTPAGACCCCGTACA	3679
Db	3718	GGTTTCCGCCCCCAAGGACG---GTACGACCTGGTGTTCATCACTTGGAACTTAAATACA	3774
Oy	3660	GGAAACCATCATACCAACAGTGGGAGGTACGCGTATCCACACAGATGTAAAGTGA	3729
Db	3775	GAACACCACTCTTTCAGCAGTGCAGAGCCATGCGGCACTTAAAAACCTTTCCGGCTT	3834
Oy	3740	AGGCTGTCCACCACTGCAACTGGCGGAATGTGTGCTATAGGGTATGGGCTTGGCTG	3799
Db	3835	CGGCCCTGAATTGGCTTAAACCCAGAGGACCCCTCGTGTGAATCTATGCTATGCAACCG	3894
Oy	3800	ATCCGGCAACCGAGAATATCATCACTGGGTGGACGCTCATTTTAGTTTACCCTGTCT	3859
Db	3895	ACCGCAACAGTAGAGACGTAGTCAACCCCTCTTCCAGAAAGTTGTCAAGGCTGTCCAG	3954
Oy	3860	GTCAAGCTTAAGAACACGTCCGCAAAATCTAGTCTCTTCTGTCTTCTTCCGCAAGACA	3919
Db	3955	CGAGACCAAGATTGTGTCTCAAGCAATACAGAAATGTACCTGATTTTCCGACACTAGCA	4014
Oy	3920	ACGGCAACCAACAATGACAGGACAGACTCGGTAGTGCCTTGACAAACATCTAT--CAA	3978
Db	4015	ACAGC--CGTACACGGCAATTCAACCCCGCACATGTGAATTCGTGATTTCCGTGTGAT	4073
Oy	3979	GGGTCAACCAAGTACGAGCGAGGAGAGCTTCAAGCTTACAGATGATCAGGTGACATT	4038
Db	4074	GAGGGTACAAAGATGAGTGGAGTGGACGCGCGCTCATACCGCACAAAGGGAATATTT	4133
Oy	4039	AGCAAGACCGCTGACCAAGCTATGTTAATGTCTTAATAGCAAAAGTGCACAGGTTCC	4098
Db	4134	GCTGCTGTCAAGAGGAGCAAGTTGTCAACGCAACCAATCCGCTGGGTAGCCAGGCGAA	4193
Oy	4099	GGAGTGTCCGCTGCACTGTACCGAAATATGCGCGCTCTTTGTATAGACAGCCATAGCT	4158
Db	4194	GGAATGTCCGCTGCATCTTAATAACGTTGGCGACAGTTTTACCGATTACACACGAG	4253
Oy	4159	GTCCGGACGGGTAGCT---TGTGAAGCAGAACCGCTCATCATATGCTGTAGACC	4215
Db	4254	ACAGGCAACCGCAAGATGACTGTGTGCTAGGAAGAAAGATGATCCACGGCGTCCCT	4313
Oy	4216	AATTTTTTAAAGATGCCGGAACCGGAGGCGCACTTAAAGCTCGGACGTCTTACATGAC	4275
Db	4314	GATTTCCCGAAGCACCCGAGAGCAAGAGCCTTGAATTTGTCAAAAACGCTTACATGCA	4373
Oy	4276	ATAGGTCATCTGTCAACGCTGAGCGGTTTCAAAAATATCAGTACCGCTACTGTCAAC	4335
Db	4374	GTGGCAGACTTAATTAATGAACATTAACATCAAGTCTGTGCCATTTCCACTGTATCTTACA	4433
Oy	4336	GGCATCTATTTCTGTGCGAAGATCGAGTGTGCAATCATTTGCATCTGTTCACTGCT	4395
Db	4434	GGCATTTTACGACGCGGAAAGACCGGCTTGAATATCACTTAACCTGTTGACAAACGCG	4493
Oy	4396	TTCCGACTACGATGCCGATGTCAACCATATATTGCTTGGATTAACAATGGGAGACAGG	4455
Db	4494	CTAGACATATCTGACGCGAGTAAACATCTATTGCTGTGATTAAGATGGACAGGAAGA	4553
Oy	4456	ATATTCGAGGCGCATTCACCGCAAGAAAGGTGCAATTTCT--GGATGATGACAGGCA	4512
Db	4554	ATCGACGGGCACTCCAACTTAAAGAGTCTGTAAACAGCTGAAGAGATGAAGATATGAG	4613
Oy	4513	GTAGACATTAAGTGTGTACGGGTCCACCAACAGCTCTTGGCAGGCAAGCCAGTTTAC	4572
Db	4614	ATCGACATATGATTAATGATGATTCATCTCAGACAGTTGCTTGAAGGGAAGAAAGGATTC	4673
Oy	4573	TCCGTCAATGAGGCAAGTTGATTTCAATCTGGAAGTACAGATTCATCAGACCGCC	4632

1037 ATCAGATGACGAGGATTCTGGCAACTGACCTTAGTGTGATGACGACCAAAAACTATTGG 1096
1081 ATGAGATGATGATGATTAATGGCCAGATATATCACTGACGATGACAAAACTTTGG 1140
1097 TTGGGCTCAACCAAGATTGTCTCAATGTAGAGCGCAAGAAATATCTAACAAATGC 1156
1141 TTGGGGCTCAACGAGCAATGTCTAATGAGTAACTAAGAACACCAACACCAATGC 1200
1157 AGAATATATCTAATACGAGGTGGCCAGGCTTTTCCAGTGGGCGCGGAACATGTG 1216
1201 AAAATTAACCTTCTGCGCATATAGCACAAAGGTTACAAATGGCTAAGAGGCAAG 1260
1217 CCGACTTGACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTACTATGAGGCTGT 1276
1261 ATATCTTGATTAACGAGAAAAATGTGGTACTAGAAACCGAAGCTTATCGTATGGCT 1320
1277 GCTGGGCTTTCAGAACCCAGAAAAATCATCTACAAAGAGCTGTATCGCAACAA 1336
1321 TGTGGGCTTTCGACATTAAGAAAGTACATTCGTTTATGCGCCACCTGGAACGACACT 1380
1337 TTAAAGAAAGTACTGCGCTCTTGAATCTATTTGATTCACAGCTTATCCAGACGAGG 1396
1381 GCGTAAAGTCCAGGCTCTTTTAAAGCGCTTTCCATGTGCTCGTATGACACACTTT 1440
1397 TCGATATGGGCTTCCGCGTAGGCTCAAGCTGCTGCTTGAACCA--ACTGTCAACCG 1453
1441 TGGCCATGTGCTGAGGCAAAATTTGAATCTGGATTGCAACCAAGAAAGAGAAAAAC 1500
1454 CACCGCTATTAACATGCGGATGTGAGACATCTGCGTGTCTTACAGCAAGAAAGCTGAG 1513
1501 TGTGCGAGGCTCTCGAGGAATTAAGTATGAGAGCCAAAGCTGCTTTGAGAGATGCTCAG 1560
1514 AATGTGCTGAGGAGAAAGATCAGAAAGCCCGCAACCTTCTCTCCGAAATGAAA 1573
1561 AGAAGCGAGGCGGAGAAAGCTCGAGAACACTTCCACATTAATGTGAGACAAAGGCA 1620
1574 AAGAGACCTGAGAGCA-----GAAATGACCTCATTTATGCAAGAGGAGAGAGCAG 1624
1621 TCGAGGAGCGCGAGAAAGTTGTCTGCGAAGTGGAGGGGCTCCAGGCGGACATCGAGCAG 1680
1625 GTAGCGTGAGACACACGAGACACATCAGGGTGACAAAGTTAACAGCGGAGAGAGAA 1684
1681 CATTAGTTGAAACCCCGCGGCTCACGTAAGATTAACCTCAAGCAAAATGACCGTATGA 1740
1685 TTGGGCTTACGCTAATCTTCAACCCAGGCGGATTAATGAGAAAACTTGGCGGTGA 1744
1741 TCGGACAGTATATGTTGTCTCGCAAACTGTGTGAAAGAAATGCCAACTGCAACAG 1800
1745 TCCACCATTTGGGGAACAAGTACTGGTAAATGACTCAAAAGGTAGGGAGAGAGATACA 1804
1801 CGGACCGCTAGCAGATCAGGTTAAGTCAATTAACACCTCCGAAAGATCAGGAAGGTACG 1860
1805 AAGTGAAGCATACACGATTAAGTCAATGTAACGAAGGAGCGCGGTCTCTGTTCAAG 1864
1861 CGGTGCAACATACGAGCTTAAGTATGATGACAGAGAGAGGTGCGTACATGAGCAG 1920
1865 ACTTCAAGGATGAGAGAGAGGCTAAGTCTGTTTCAACGAGAGGAAATTTGTTAAACA 1924
1921 AATTTCTAGACATGAGTGAAGGCGCACTTAAGTACAAAGAAAGAAATTTGTGAACC 1980
1925 GATACCTGACACATCGCATCAACGAGAGAGGCTTAACACTGACGAGAGATACATATA 1984
1981 GCAATCATATACCATTTGCTCATGATGCCCCCGCAAGAAATACGAAGAGAGACATACA 2040
1985 AGACTGTAAGACTCAGAGACAGACTCAGAAATACGTTCTTGAATTTGACGCAAGAAAGT 2044
2041 AGGTTCAAAAGGAGAGCTTGAGAAACAGATACGTTTGAAGTGAAGCAAGAGGCTT 2100
2045 GTGTTAAGCAGAAAGCAGAGTCCCTTGCGCTTAACCGGTGATCTGTGATCTCAACAT 2104
2101 GCGTTTAAAGAAAGAAAGCTCAGGTGTGTTCTCTGAGGAATCTAACCAACCTCTCT 2160
2105 TTCAAGGTTTGGTACGAGAGTCTCAAGACAGACAGACAGACACTCAAAAGTCCAA 2164

2161 ATCATGAGGTAGCTGTGAGAGGACTGAGACCCGACCTGCGGTCCTGTAAGAGTCAAAA 2220
2165 CCATCGAGCTCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGCTGTGA 2224
2221 CATTAAGAGTGAATGAGCAACCGGGGTGGGCAAGTACACTAATTAATCAAGTCACTGTCA 2280
2225 CTAAAGAAAGTCTGTTGTGAGTGCAGAGAAAGAAATCGCGCAAAATCATCAGGATG 2284
2281 CGGACGAGATCTTTGTATACGCGGAAAGAAAGAAATTTGTCCGAAATTTGAGCCGACG 2340
2285 TAAAGAGATGACGATATGATGTGCTGTAAGACATGCAATGAGTCTTCTAAATG 2344
2341 TGTAAAGACTGAGGGGTATGACAGATTACGTGCAAGACAGTAAATGCGTTATCTCAACG 2400
2345 GGGTTAAGCAACCCGTTAACAATCTGTATACATTAAGGCAATTTGCTGCTGATGACGGA 2404
2401 GATGCCACAAAGCCGTAGAAAGTGTGATGATGACAAAGGTTGCGGTGACAGCAGAG 2460
2405 CGGTGCTGGCACTGATTCGATGTCAAAAC---TAAGAAAGTGTATTGTGGGGGAGC 2461
2461 CACTACTGTCTGATTTGATGCTATGTGAGGCCCCGCAAGAGTATGATCTATGCGGAGACC 2520
2462 CAATAATGCGGCTTCTTTAATGATGATGCTGAAAGTACATTTAAAC----- 2512
2521 CCATGACATGCGGATCTTCAACATGATGCACTAAAGGATACATTCACCTGAAA 2580
2513 ATGACATATGCACTGAAGTGTACATTAAGAAAGCTCTGTAGAGGTGACACAGACTGTAA 2572
2581 AAGACATATGACACCAAGACATCTACAAATATATTCGCGGCTTTCACACACACAGTTA 2640
2573 CCGCATGCTGTACAGCTCTTCTACAGCAAGGAAAGAAAGGTTAAACCATGTCG 2632
2641 CAGCTATTGTATGACATCTGATTAAGATGAAAGTGAAGAACCAAGAACCTGCAAGA 2700
2633 ATAAATCATATATGATTAACCAACAGGAGACCAAGGCGGACAAAGATGATCTGATTTAA 2692
2701 AGAAATTTGAATTCGATATTAACAGGGGCCCAAGGCGCAAGGAGATATCATCTGA 2760
2693 CCTGTTTCAAGAGATGGGTGAAGACAGCTACAGATTAATGATCAAAATTCAGAAATCATGA 2752
2761 CATGTTTCGCGGGGTGGTTAAGCAATTCAAATGACATATCCGGAATGAAGTATGA 2820
2753 CTGCGGCTGATGCAAGGACTTACGCGGAAAGGCGTTATGCTGACAGTAAAGTCA 2812
2821 CAGCGCGGCTTACAAAGGCTTACCAAGAAAGAGTGTATGCGCTCCGCAAAAAGTCA 2880
2813 ACGAGATCCACTCTACTGCAAGCTTCTGAGCAGCTGAACGTGTTACTTACAGCAGCAG 2872
2881 ATGAAACCCACTGTACGCGATCACAATCAGAGATGAAACGTGTGCTCACCGCACTG 2940
2873 AAAAAGCATTTGCTGGAAGACGCTAGCTGTATCTCTGATTAAGACATTTACAGCTA 2932
2941 AGGACAGGCTAGTGTGAAGAACTTGCAGGCGGACCCCATGAGATTAAGACAGCCCATACA 3000
2933 AATATCCGCGGATTTCAAGGCTTCAATGAGACACTGAGCGCGCAACAGCGCATTA 2992
3001 TACTTAAAGGAATCTTACAGGCTACTATAGAGACTGGAAGCTGAACACAGGAAATA 3060
2993 TGGCAGCGTCTTGTATGAAGCCGACAGACTGATGTGTTCCAGAAATGAGTAAAGTCT 3052
3061 TTGCTCAATTAACAGACCCCACTCCCGTGCATTCGTTACAGCTCAAGACCAAGCTTT 3120
3053 GCTGGCGGAAGGCTTTAGAGCCAGTCTTGCCCAACGCGCAACATTTGTCTGACAGACAGC 3112
3121 GCTGGCGGAAGGATTTGAACCGATCTAGCCACGCGGTATCGTATCCGCTGTC 3180
3113 AGTG---GAGAGCTTGAACCCATTAAGCATGACAGAGGTCTCACTGAAATGAGCAC 3169
3181 AGTGAGCGAATCTTTTCCACAGTTTGCAGATGACAAACCAATTTGGCATTTAAGCTC 3240
3170 TGAATCTTTTTCACAGAGTCTTTTGAATGATCTGACAGTGGGTTATTTTCGCTC 3229

Db 3241 TAGAGTAATTTGCAATTAAGTTTTCGCGATGGAATGCAAGCGGACTGTTTCTTAAC 3300
Qy 3230 CTACCGTCGCACTTACTTACAG-----GGATGACGACTGGGATA 3268
Db 3301 AGAGCATCCCACTTAAGTACCATCCCGGATTTACGCGAGCCGGTAGCTCATTTGGGACA 3360
Qy 3269 ACTCGCGAGGAGAAATGATGATGGCTTAATAGAGGTAGCAAGAGTTGTCAAGC 3328
Db 3361 ACAGCCGAGGAGAACCGCAAGTATGGGTACGATCAGCCCATTTGCCCGCAACTCTCCGTA 3420
Qy 3329 GATATCCGTCATCAACAAAGCGGTGACACAGCGAGGTAGCTGATTAATAGGAATATA 3388
Db 3421 GATTTCCGGTGTTCAGCTAGC---TGGGAGGGGACACACACTTGTATTTGACAGCGGGA 3477
Qy 3389 CATCAAGGACTACTCTCCCAATTAATGATGTTTCATTAATAGCGCGGTGGCCCACT 3448
Db 3478 GAACAGAGTATCTCTGACAGCATTAACCTGTTCCGGTGAACGCAATCTTCTCAGC 3537
Qy 3449 CGTTATCGTTGACCAAAAGGACAGGGTACATGATCAGCGGATTCATCTAAGA 3508
Db 3538 CTTATGTCCTCGAGTACAAAGAGCAACCGGCGGCTCAAAAATTTCTGAACCACT 3597
Qy 3509 TGAAGGCAATCTGTGTGTGTATC---GGGATCTATCAGCATTCAGGAAAGAAAG 3565
Db 3598 TCMAACACCACTCAGTACTGTGTATCAGAGGAAAAATTAAGCTCCCGTAAAGAA 3657
Qy 3566 TAGAGTCAATGGGTCCATTTGCCAC-----TAATACATCAGGTGTATCTCGATTTGG 3619
Db 3658 TCGAATGATCGCCCGATTTGGCATAGCCGGTGCAGATAGACTAACAATCTGGCTTTCG 3717
Qy 3620 GAATACCTAGCCATGCGTAAATATGACATTAATCTTTGTCATGTTAGAACCCCGTACA 3679
Db 3718 GGTTCGCGCGAGGAG---GTACGACTGTGTATCAACATTTGAATTAATACA 3774
Qy 3680 GGAACATCATCAACAAGTGCAGGATCAAGCTATCAACAAGCATCTTAACGTGTGA 3739
Db 3775 GAAACACCACTTTCAGCAGTGCAGAACCATAGCGGCACTTAAAAACCTTTGGCGGT 3834
Qy 3740 AGGCTGTCCACCACTGAAACATGGGCGGAACTGTGTGCTATAGGATATGGGCTTTCG 3799
Db 3835 CGGCTTGAATTCCTTAACCAAGAGGACCTGTGTGAATCTAATGCTTACGCG 3894
Qy 3800 ATCGGCAACCGAGATATCATCACTGCGGTGACGCTCATTTAGGTTTACCGGTCT 3859
Db 3895 ACCGAAACGTGAGAGTACGCGCTCTTCCGAAAGTTTGTACAGGGTGTCTGAG 3954
Qy 3860 GTACCTTAAGAACCTGCGGAAATTAAGAGTTCTTGTGTCTTTCGCAAGACA 3919
Db 3955 CGAGACCAAGATTGTCTCAAGCAATACAGAAATGATCACTGATTTCCGACAACTAGACA 4014
Qy 3920 ACGGCAACCAACATATACCAAGACAGACTGGGTATGTTGATCAACATCTAT-CA 3978
Db 4015 ACAGC-CGTACCGGCAATTCACCCGCACTGTGAATTTGGTATTTTCCGCTGAT 4073
Qy 3979 GGGTCAACCGAGTACGAGCAGGAGAGCTCCAGCTACAGATGATCAGAGTACAT 4038
Db 4074 GAGGCTACAAAGAGTGAAGTTGAGCCGCGCTCATACCGAACCAAAAGGAGAAATTT 4133
Qy 4039 AGCAAGAGCGGTACCAAGCTATCGTTAATGTCTTAATAGAAAAGTCAACAGTTCC 4098
Db 4134 GCTGCTGTCAAGAGAAAGTGTTCACAGCAATCCGCTGGTATAGACAGCGCA 4193
Qy 4099 GGAGTGTGGGTGACCTGTACGGAATAGCCGGTGTCTTTGATAGACAGCAATAGCT 4158
Db 4194 GGAGTGTGGGTGACCTGTATTAAGCTGGCGACCAAGTTTACGATTTACGCGAG 4253
Qy 4159 GTCGGAGCGGTAGACT---TGTGAAGCAAGAACGCTCATCATCATGCTTAGAGCC 4215
Db 4254 ACAGGCAACCGAGATAGCTGTGTGCTTAGAGAAAGATATCAACCGGTCGCGCT 4313
Qy 4216 AATTTTCTAAGATCCCGAAGCGGAGCGGACCTTAAGCTGCAAGCTGCTCATGAGC 4275
Db 4314 GATTTCCGGAAGCAACCGAAGCAGAGCTTGAATTTGTAACAAAGCCTTACATGCA 4373

Qy 4276 ATAGGTCATGTCACAGCTGAGGAGTTACAAAATATAGTACCGCTACTGTCAAC 4335
Db 4374 GTGCGAGACTTATGTAATGAACTAATCAAGTGTGTGCGATTCACATGTAATCA 4433
Qy 4336 GGCATCTATTCTGTGGCAAAAGATGAGTATGCAATCATTTGACCACTGTTCACTGT 4395
Db 4434 GGCATTTACGAGCGGAAAGAACCGCTTGAAGTATCACTTAATGTGTTGCAACCGG 4493
Qy 4396 TTGCACTAGGAGATGCGGATGTCAACATATTTCTTGATTAACAAATGGAGACAG 4455
Db 4494 CTAGCAAAACTGACGCGGAGTAACTCATTTCTCGATTAAGAGTGAAGGAAAGA 4553
Qy 4456 ATAACTGAGGCACTACCGGAAAGAACGCGAAATTT---GGATGATGACAGCA 4512
Db 4554 ATCGACGCGGCACTCAACTTAAGAGTCTTAACAGAGCTGAAGAGTAAATATGAG 4613
Qy 4513 GTAGCATTTGACTTGTGAGGATCCACCAACAGCTTTTGGCAGGACACAGGTTAC 4572
Db 4614 ATCGAGATGATGATGATGATGATCCAGAGTGTCTTGAAGGAAAGAGGATTC 4673
Qy 4573 TCCGTCATGAGGCAAGTTGATTTATCACTGCAAGTACAGATTCATCAGACCGC 4632
Db 4674 AGTACTACAAAGGAAATGATTTGATCTTGAAGGACACCAATTCATCAACAGCA 4733
Qy 4633 AAGCATTTGCGGAATCCATGCAATGTGGCCCAACAAATGAGGCTAATGAGAGATT 4692
Db 4734 AAGCATGAGGAGATTAAGGTCTGTTCCCTTAATGACAGGAAATGAATGAACATG 4793
Qy 4693 TGTCTGATCATCTGTGGGAGAGTATGTCCAGCATCCGCTCCAATGCGAGTAGAGAG 4752
Db 4794 TGTGCTATCATATTTGGTGAAGACATGGAACAAATCCGCAAAAGTGTCCGCTGACAT 4853
Qy 4753 TCAAGGCGGTGCTCTCACTTCAACACTTCATGCTGTGTATTAATGCTATGACGCT 4812
Db 4854 AACCGTGTGAGCCCGCAAAAGCTGTGCTTGCATGATGATGATGATGATGATGATG 4913
Qy 4813 GAGGCGGTATCAGTTGCGCTGTGCGAAGAAAGAGTTCGCGTATGCTCATCATTC 4872
Db 4914 GAAAGGTCACAGCTTAAGAAAGATTAAGTCAAGTAAAGTAAAGTAAAGTAAAGTAA 4973
Qy 4873 CTGTTCCGAAGTACAGATCAAGCGGTGAGAAAGTATACAGTGCAGCAACAGTCTG 4932
Db 4974 CCCCTCTTAAGCAAAATTAAGAAATTTAAGAAATTTAAGAAATTTAAGAAATTTAAG 5033
Qy 4933 TTTTCAGGCGTGTACACCGGCTGTACACCCAGAAAGTACGCGGAAT 4982
Db 5034 TTTAATCCGACACTCCCGCATTCGTTCCCGCGTAAATGATGAGAT 5083

RESULT 12
US-09-507-362-101
; Sequence 101, Application US/09507362
; Publication No. US20030096397A1
GENERAL INFORMATION:
APPLICANT: Dubensky Jr., Thomas W.
; Polo, John W.
; Belli, Barbara A.
; Schlesinger, Sandra
; Dryga, Sergey A.
; Prolov, Ilya
TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
SYNTHESIS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESS: Seed Intellectual Property Law Group PLLC
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,362
FILING DATE: 18-Feb-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 930049.457D6 /1196.011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 8000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 101:
US-09-507-362-101
Query Match 25.9% Score 1295.6; DB 11; Length 8000;
Best Local Similarity 56.0%; Pred. No. 0;
Matches 2816; Conservative 0; Mismatches 2134; Indels 80; Gaps 16;
QY 23 ATATGAAAGAAATTCACGTTGACTTAGTGAAGCCGCTATGCTCAAGTCGTTACAGC 82
DB 64 AGAAGCCAGTAGTAACGTAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
QY 83 GGAGCTTCCCAATTTGATGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 142
DB 124 AAACCTTCCCAATTTGATGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183
QY 143 CCAGAGCGTTTGGATGAGCAAAAGCTCATTTAGAGAGAGAGAGAGAGAGAGAGAGAG 202
DB 184 CCAAGGCTTTTGGATGAGCAAAAGCTCATTTAGAGAGAGAGAGAGAGAGAGAGAGAG 243
QY 203 TTAATCTTGACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 262
DB 244 CGATCTTGACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 303
QY 263 GTATCTGCGCTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 322
DB 304 GTGCTGCGCTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
QY 323 TTAAGAAAAA-----GTGACATTTACCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 376
DB 364 TGGCGGAAAAAGCGTGCAAGATTTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
QY 377 TGGAGATGATGCAACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
DB 424 GGACGATCTTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
QY 437 CGTGTAGTACTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
DB 484 CCGTGAACATCGTGGCCGAATATTTCCGTATGACAGAGAGAGAGAGAGAGAGAGAG 540
QY 497 CATCAATGTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556
DB 541 GAACTATCTATCATAGGCTATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 557 CGAGCCCTTTTATGATGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
DB 601 CCAACCAATGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 617 CTGACGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 676
DB 661 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 677 GCAGGCTTGAAGAACTCTCAATCTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736

DB 721 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 737 TATTTCTGGTTGGTTCAACATCTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 796
DB 781 ATTTCTCGTAGAGATCCACATTTATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 797 TTCCAAACGTTTCCATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 856
DB 841 TTCCATCGGTGTTCCATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 857 TCAAGCTGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 916
DB 901 TGAAGTCGAAGGCTAGCTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 917 TTGAGAACTTGGGCTCAACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 976
DB 961 CCGTGGAGATACGGGTTCACACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 977 CGCTGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1036
DB 1021 CAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1037 ATCAGATGACAGAGATTTGAGCAATGACGTTAGTGAAGAGAGAGAGAGAGAGAGAGAG 1096
DB 1081 ATCAGATGACGAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1097 TTGGGCTCAACCAAGAGATTTGCTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1156
DB 1141 TTGGGCTCAACCAAGAGATTTGCTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1157 AGAAGCTATCTATTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1216
DB 1201 AAATTTACCTTCTGCGATCATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1217 CCGACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1276
DB 1261 ATGATCTTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1277 GCTGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1336
DB 1321 TGCTGGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1337 TTAAGAAAGTACCGGCGCTTTGACTATTGATTTCCAGCGCTTACAGAGAGAGAGAG 1396
DB 1381 GCGTAAAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1397 TCGATATGAGGCTTCCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1453
DB 1441 TGCCCATGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1454 CACCGGCTATTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
DB 1501 TGCTGAGAGTCTCGAGAGAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1514 AAGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1573
DB 1561 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1574 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1624
DB 1621 TCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1625 GTAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1684
DB 1681 CATTAAGTTGAAACCCCGCGCTGACGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1685 TTGGGCTTACGATATCTTCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1744
DB 1741 TCGAGAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1745 TCCAGCAGATTTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1804

Db 1801 CGCACCCGCTAGACAGATCAGGTTAAGATCATTAACACTCCGGAGATCAGGAAGGTACG 1860
Qy 1805 AAGTCGACCCATATACACGAGTAAAGTCAATTGTATACAGAAAGGACGGCGGCTCCCTGTTCAAG 1864
Db 1861 CGGTGAAACCATACGACCGCTTAAAGTACTGATCCACGACGAGAGTCCCTGATCCATGCGCAG 1920
Qy 1865 ACTTCGAGGCAATTGAGTGAAGACCGGTACGATGTTTCAACGAGAGGAGTTCGTAAACA 1924
Db 1921 AATTCCTGACACTGAGTGAAGCGCCAGCTTAGTATACACGAAAGAGGCTGTGAAACC 1980
Qy 1925 GATACCTGCACCACTCCGCAATCAACGAGAGGCGCTAAACCTGACGAAAGATCTATA 1984
Db 1981 GCAAACTATACCAATTCGCAATGACGCGCCGCAAGATACAAAGAGGAGCGTACACA 2040
Qy 1985 AGACTGTAAAGACTCAGACACAGACTGAGAAATACGTCCTTGCATTTGACGACGAAAGT 2044
Db 2041 AGGTTACAAAGGACAGCTTGACGAAACAGAGTACGTGTTGACCTGACGAAAGACGTT 2100
Qy 2045 GTGTTAAGCGAAGACGCAAGTCCCTTGTCCTAACCGGTGATCTGTAGATCCACAT 2104
Db 2101 GCGTTAAGAGGAAAGACCTCAGGTCGTGTCCTCCGGAGAACTGACCAACCTCCCT 2160
Qy 2105 TTCAGAGATTGGCGTACGAGAGTCTCAAGACACGACGACGACCTCAAAAGTCCCA 2164
Db 2161 ATCATGACCTAGCTCTGAGAGGACCTGAAGACCCGACCTGCGGTCCCGTACAAAGTCCGAAA 2220
Qy 2165 CCATCGAGTCTATGAGAGTCCAGGTTCAAGTAACTGSAATCATCAAAAGCGCTGGA 2224
Db 2221 CATATGAGATGATGACACACCGGGGTGGGCAATCAGCTATTATCAAGTCACTGTCA 2280
Qy 2225 CTAAAGAAAGATCTGTTGTGAGTGCAGAAAGAAAGAACTGCGCAGAAATCATCAGGGATG 2284
Db 2281 CGGACGAGATCTTGTTCACAGCGGAAAGAAAGAAATTTGTCGCAAAATGAGGCGCAG 2340
Qy 2285 TAAAGAGATGAGACGATGATGTTGCTGCTAGACCTGTGATTCAGTCTTAAATG 2344
Db 2341 TGTAAAGCTAGAGGGGTATGCGATTAGCTGAAACAGTAGATTCCGTTATGCTCAAG 2400
Qy 2345 GGGTTAAGCAGCCGCTTAAACCTGTAATTGATGAGAGCAATTCCTGCGCATGAGGGA 2404
Db 2401 GATGCCAAGAGCGGTAGAGAGTCTGTACGTTGACGAAAGCGTCCGTCGCAACGAGAG 2460
Qy 2405 CGCTGCTGCACTGATTCGCACTGCTCAAAAC---TAAAGAAAGTGTATTGTGCGGGAGC 2461
Db 2461 CACTACTGCTGATGTTGCTATCGTCAGGCCCCGCAAGAGTACTATGTCGAGAC 2520
Qy 2462 CAAACCAATGCGGCTTTTAAATGATGTCGTAAGATCATTTTAAAC----- 2512
Db 2521 CCATGCAATGCGGATCTTCAACATGATGCACTAAAGTACATTTCAATCACCTGAAA 2580
Qy 2513 ATGACATATGACATGATGATACCATTAAGCATCTTAGAGGTGCACAGACTGTAA 2572
Db 2581 AAGATATATGACCAAGACATTTCTAACAGTATATCTCCGGGTGTCACACAGCCAGTTA 2640
Qy 2573 CCGGCATGCTGTCACGCTTCTCTACGACAAAGCATGAAAGAGCGGTAAACCATGTGCTG 2632
Db 2641 CAGCATATGATTCGACATGCTTACGATGAGAAATGAAACACAGAACCCGTGCAAGA 2700
Qy 2633 ATAAATCATATAGATACCAAGGAGACCAAGAGCCGCAAAAGATATCTGATTTCA 2692
Db 2701 AGAATATGAAATCATATTTACAGGGGACCAAAAGCCGAGGAGATATCATCTGGA 2760
Qy 2693 CCTGTTTCAAGAGATGAGTGAAGACGCTACAGATTGATTAACAAAATACGAAATCATAGA 2752
Db 2761 CATGTTTCCGGGGGTGATTAAGCAATTGCAATGACATATCCCGAGATGAAAGTAAAGA 2820
Qy 2753 CTGCGGCTGCAATCGCAAGACTTACGCGAAAGCGTTTATGCTTACAGTAAAGTCA 2812
Db 2821 CAGCGCGGCTCTCAAGAGGCTTACAGAAAGAGTGTATCCCTCCGGCAAAAGTCA 2880
Qy 2813 ACAGATATCACTTACTGACAGACTTGTAGACAGTGAACGTGTTACTTACACGACAG 2872
Db 2881 ATGAAACCCACTGATACCGCATCATCATGAGCATGTGAACGTGTTGCTCACCCGACACTG 2940

Qy 2873 AAAACGCAATGCTGGAAGACGCTAGCTGTGATCCCTGATTAAGACACTTACAGCTA 2932
Db 2941 AGGACAGGCTAGTGTGAAAAACCTTGACGGCGACACCATGATTAAGACAGCTCACTACA 3000
Qy 2933 AATATCCCGGGGATTTCAAGGCTTCATTTGACAGCTGGACCGGCAACAGACGCATTA 2992
Db 3001 TACCTAAAGAACTTACAGGCTACTATAGAGAGCTGGGAAGCTGAAACAAAGGAAATTA 3060
Qy 2993 TGGCAGCGTCTTGATTAAGCCGACAGACGTGATGTGTTCCAGAAATAGGTGAACGCT 3052
Db 3061 TTGCTGCAATTAAGACCCCACTCCCGTGCATTCGTTACGCTGCAAGACCAAGCTTT 3120
Qy 3053 GCTGGGCAAGGCTTTAGAGCCAGTCTGGCGCACGGCAACATTTGTGCTGACAGACAGC 3112
Db 3121 GCTGGGCAAGGCAATGGAACCGATCTAGCCACCGCGGATGCTTACTTACCGTTGCC 3180
Qy 3113 AGTGG---GAGACGTTGCAACCATTTCAAGATGACAGAGCTACTCACCTGAATGGAC 3169
Db 3181 AGTGGAGGAATGTTCCCAAGTTGGGATGACAAACCAATTCGCGCATTTACGCT 3240
Qy 3170 TGAATCTTTTGAACAGGTTCTTTGAGATGACCTGGAACAGTGGGTTATTTCCGCTC 3229
Db 3241 TAGACGTAAATTTGATTAAGTTTTCGCGCATGACCTTGACAAAGCGGACTGTTTCTAAAC 3300
Qy 3230 CTACCGTGCACCTTACTTACAG-----GGATCAGACACTGGGATA 3268
Db 3301 AGAGCATCCCACTAACGATACCATCCCGCATTTACAGAGCGCGGTAGCTATTTGGACA 3360
Qy 3269 ACTGCCAGGAGAAACATGATGCGCTTAATAGAGAGTGAACAAAGAGTGTTCACGCG 3328
Db 3361 ACAGCCAGGAACCGGAATGATGGGTACATCAAGCCATTTCCGCGCACTTCGCTGA 3420
Qy 3329 GATATCGTGCATCAAAAGCGGTTGACACAGGCGGTGCTGATTAAGAAATATA 3388
Db 3421 GATTTCCGCTGTTCCAGCTAC---TGGGAAGGCAACAACCTGATTTGACAGCGGGA 3477
Qy 3389 CCATCAAGAGACTACTCTCCCAATTAATGATTCATTAATGCGCGGTGCCCCACT 3448
Db 3478 GAAACAGATTAATCTCTGACACAGATTAACCTGATCCGCTAACCGCATTTCTTCACG 3537
Qy 3449 CGTTGATGTTGACACCAAGAGACAGGATCAACATGATCAAGCGGATTCCTATCTAAGA 3508
Db 3538 CTTATGTCGCCGAGATCAAGAGAAAGCAACCGCGCGGTGAAAAATTTCTTGAACGAGT 3597
Qy 3509 TGAAGGGCAATCTGTGTGTGATC---GCCGATCTATACGATTCAGGAAAGAAAG 3565
Db 3598 TCAAAACACACTCAGTACTTGTGATCAGAGAAAAATGAAAGCTCCCGTAAAGAA 3657
Qy 3566 TAGAGTCCATGGGTCCATGCCCC---TAATACATGAGGTGATCCGATTTGG 3619
Db 3658 TCGAATGATGCCCCCATGCGATGAGCGGTGACATAGAACTACAACTCGGCTTCG 3717
Qy 3620 GAATATCAGGCATGTCGGTAAATATGACATTAATCTTGTGCAATGTTAGAACCCCGTACA 3679
Db 3718 GGTTCGCGCGCAGGACG---GTACAGCTCGGTGTTATCAACATTTGGAATTAATTA 3774
Qy 3680 GGAACCATCACTACCAACAGTGCAGAGATCAAGCTATTCACCAACAGATGCTTAACGTGA 3739
Db 3775 GAAACCAACACTTTCAGAGATGCGAAGACATGCGGCGACCTTAAAAACCTTTGCGCTT 3834
Qy 3740 AGGCTGTCACACACTGAACACTGGCGGAATGATGTCATTAAGGATTAAGGCTGTG 3799
Db 3835 CGGCTCTGAATTTGCTTATCTAGAGAGCACTCTGATGTAAGTCTTATGCTTACGCG 3894
Qy 3800 ATTCGCGCAACGAGAAATATCATCTGCGGTGAGACGCTCATTTAAGGTTTACCGGTGCT 3859
Db 3895 ACCGCAACAGTGAAGAGTACTACCGCTCTTGCAAGAAATTTGTCAAGGTTGCTGAG 3954
Qy 3860 GTACGCTTAAGAAACATGCGCAAAATACTAGGTTCTTGTGTTCTTCCGCAAGACA 3919
Db 3955 CGAAGCAGATTTGTCTCAAGCAATACAGAAATGTAATGATTTTCGACCAACTAGACA 4014

OY	727	TATTCGCGGTGGTTCACAATCTACACGAAGATAGATCACTGTTACTACTGGCATC	796
Db	1673	ATTTCCTCGTAGAATCGACACTTTATCCGAACACAGACCAGCTTGACAGACTGGCAATC	1733
OY	797	TTCCAAACGCTGTTCACACTTGAAGGAAACTCTAACCTTCACAGGTAGATGCGGACCATTCG	856
Db	1723	TTCCATTCGGTGTTCACCTTGAAATGGAAAGCAGTCGACCTTGCCGCTGTGATACGTG	1793
OY	857	TCAGCTGTGAAGGGTACGTCATCAAAAAGATACATCAGCCCAAGACTATAACGTTAAG	916
Db	1793	TGAGTTGGCAAGGCTACGTAGTGAAGAAAATCACCATCACTGTCGCCGGATCACGGGAGAA	1855
OY	917	TTGAGAACCTTGGCGTCCACAAATGCATCCGAGGGTTCCTTGAGTTGCCAAATGCACATAT	976
Db	1853	CCGTGGGATACCCGGTTACACAACAATAGGAGGGCTTCCTTGATATGCAAAAGTTACTGCAC	1911
OY	977	CGCTCGCGGCGAGAGGGGTTCTTTTGCCTGTGTATCGATATGATACAGCCACACATTTGCG	1033
Db	1913	CAGTAAAAAGGAGAACGGGTATGTTCCCTGTGTGACGTACATCCCGGCAACCATATGGG	1977
OY	1037	ATCAGATGACAGGGATTTCTGGCAATGACGTTATGTTGATGATACGCACAAAACATTATGG	1099
Db	1973	ATCAGATGCTGTATTAATGCGACAGGATATATCACTGATGACAAAACATCTTCG	2033
OY	1097	TTGGGCTCAACCAAGGATGTTCGTCAATGTATGAGACGCAAGAAATACTAACATATGC	1155
Db	2033	TTGGGCTCAACAGCGAATTTGCAATTAACGTTAGACTTAACAGGAACCAACCAATGC	2097
OY	1157	AGAACTATCTATTACAGTGTGCGCCAGGCGCTTTCCAGTGTGGCGCCTGAACATCGTG	1211
Db	2093	AAATATTACTTTGCGGATCATAGCACAAAGGTTACGAAATGGCTTAAGGCGCCAMGG	2157
OY	1217	CCGACTTGGACGACGAAAGAACTAGGGGTCCGGAGCGCACTTTACTATGGGCTGCT	1276
Db	2153	ATGATCTTGATATACGAGAAATAGCTGGGTATCAGAGAACGCAAGTTATAGTATGGCTGT	2211
OY	1277	GCTGGGCTTTCAAGACCCAGAAATCAATCCATCTACAGAGCCTGTGATCGCAACAA	1336
Db	2213	TGTGGGCTTTGCGCACTAAGAAAGTACATTCGTTTATCGCCCACTGGAAAGCGACAGCT	2272
OY	1337	TTAAGAAAGTACCTGCGCTCTTGACTATTTGTATTCACGCGCTTACCAACCCAGCGGC	1396
Db	2273	GCGTAAAGTCCAGCCTCTTTAGGGCTTTTCCATGTGCTCGTATGAGAGACCTCTT	2333
OY	1397	TGCATATGGGCTTCCGCGTAGGCTCAAGCTGCTCTTGAACCA--ACTGTCAAACCG	1453
Db	2333	TGCCATATGTCCGTGAGGCGAAGAAATTGAACTGTGCATGTCAACCAAGAAAGAGGAAAA	2397
OY	1454	CACCGCTATTACAAATGCGCGATGTGAGCATCTCGTGGCTTACAGCAAGAAAGCTGAAG	1513
Db	2393	TGCTTCAGAGTCTCGGAGGAATTAATGATGAGAGCCAAAGGCTGCTTTGAGGATGCTCAGG	2455
OY	1514	AAGTGTGTGACCGGGAAGATACAGAAAGCCCTGCAACCTTGCTCCTGAATATGAAA	1573
Db	2453	AGGAAGCCAGAGCCGAGAAAGCTCCGAGAAAGCACTTTCACATTAATGTGCAGCAAAAGCA	2512
OY	1574	AAGAGACCGTAGAGGCA-----GAAGTAGACCTCATTTATGCAAGGAGGACAGAGGAG	1622
Db	2513	TCGAGGACCGCGCAGAAGTGTCTGCGAAGTGAAGGGGCTCCAGCGGACATCGAGAGCAG	2577
OY	1625	GTAACGTGAGACACACACAGGACACATCAGGGTACAAATTAACCAAGCGCAAGAGAAG	1684
Db	2573	CATTAGTTGAACCCCGCGGCTACGTAAAGATTAATACCTCAACCAATGACCGTATGA	2633
OY	1685	TTGGGTCTTAAGCTATACTTTCAACCCAGGCGGTATTGAATGTGAAAAATCGCGGTGTA	1744
Db	2633	TCGGACAGTATATCGTGTCTCGCCAAACTGTGTCTAAGAAATCCAAACCTCGACCCAG	2697
OY	1745	TCACCCATTGGCGGAAACAAGTACTGTATATATCAACAAGAGTGTGGGACAGGAGATACA	1804
Db	2693	CGACCCCGCTACGATACAGTTAAGATCATTAACAACCTCCGAAAGATCAAGAAAGTACG	2755
OY	1805	AAGTCAGACCAATCACGGAAGTCAATTTATACAGAAAGGACGGCGGTTCCCTGTTCAAG	1864

Db	2753	CGGTGAACCAATGACCCCTAAAGTACTGATGCCAGCAGGAGGTCCGTACCATGGCCAG	2812
Qy	1865	ACTTCACAGCAATGAGTGAAGAGCCCTACATCGTTTTCAACGAGAGGGAGTTCTTAACA	1924
Db	2813	AATTCTTAGCACTGAGTGAGAGCCGCACTGTAAGTGTACACGAAAGAGAGTTGTGAACC	2872
Qy	1925	GATACCTGCACCACTCCCAATCAACGAGGAGCGCTAAACACTGACGAGAGATCTATA	1984
Db	2873	GCAAACTTATCACCACTTCCCATGCAATGCCCCCGCAAGAAATACGAAAGAGCAGTACA	2932
Qy	1985	AGACTGTAAAGACTCAGGACACAGCTCAGATACGTCTTGATATGACCCAGAAAGT	2044
Db	2933	AGGTTTACAAAGCAGAGCTTGACAAACACAGATACGTGTTTGACGTGACAAAGCGTT	2992
Qy	2045	GTGTTAAGCAGAAAGACCCAGAGTCCCTTGTCCTCAACCCGATCTGGTAGATCACCAT	2104
Db	2993	CGGTTAAGAGAGAAAGAACCTCCAGTCTGTCTCTCTCGGAGAACTGACCACTCTCCT	3052
Qy	2105	TTCAAGACTTTCGCTACGAGAGTCTCAAGACGACGACGACCTCACAAATGCCAA	2164
Db	3053	ATCATGACCTAGCTCTGGAGGGACTGAAGACCCGACCTGGCGTCCCGTACAAAGTCCAAA	3112
Qy	2165	CCATCGAGTCTATGGAAGTCCAGGTTCAAGTAAATCTGGAATATCAAAAGCCCTGTGA	2224
Db	3113	CAATAGAGTATATAGGCAACCCGGGCTCGGCGCAAGTCAAGCTATATCAAGTCACTGTCA	3172
Qy	2225	CTAAGAAAGATCTGGTTGTGAGTGCAGAAAGAAACCTGGCGCAAAATCATCAGGAGTG	2284
Db	3173	CGGCAACGAGATCTTGTATCCAGCCGAAAGAAAGAAATTTGTCCGAAATTTGAGGCCAGAC	3232
Qy	2285	TAAGAGAGATGAGACGATGGAATGTCCTGCTAGAGCTGTGATTCAGTGTCTTAAATG	2344
Db	3233	TGCTTAAGACTGAGGGGTATGCAATTAAGTGGAAGACAGTAGATTCGGTTATGCTCAACG	3292
Qy	2345	GGGTTAAGCAACCCCGTTAACACTCTGTACATTTGATGAGCAATTTGCTCCCTATGACAGGA	2404
Db	3293	GATGCCAATAAGCCGTGAAAGTGTGTGACGTGACGAAGCGTTGCGCTGCCACGACAGAG	3352
Qy	2405	CGTGCTGGCACTGATTTGCCATGCTCAACCC---TAAGAAAGGTGATTTGGCGGGAGCC	2461
Db	3353	CACACTTGTGCTTATTTGCTATCGTACAGCCGCCGCAAGAAAGTATGTAATATGCGGAGAC	3412
Qy	2462	CAAAACAATGCGGCTCTTTTAAACATGATGTCGCTGAAGATACATTTTAAACC-----	2512
Db	3413	CAATGCAATGCGGATCTTTCAACATGATGCACTAAAGTATCATTTCATCACCCTGAAA	3472
Qy	2513	ATGACATATGCACTGAAGTGTACCATTAAGACATCTTAGAGGTGACACAGACTGTAA	2572
Db	3473	AAGACATATGCAACCAACATTTCAACAAGTATATCTCCCGCGTTGCAACAGCAAGTATA	3532
Qy	2573	CCGCACTGTGTCACCCCTCTTCTACGACAAAGCAATGMAAGCGGTTAAACCCATGTGCTG	2632
Db	3533	CAGCTATTGTATGCAACTGCACTTATGCAATGAAAGATGAAACCAACGACCCGTGCAAGA	3592
Qy	2633	ATTAATATCATATGATATACCAAGAGACCAAAAGCCGACAAAGATGATCTGATTTCTAA	2692
Db	3593	AGAACATTTGAATGATATTTACAGGGGCCCAAAAGCCGAAAGCCAGGGGATATCATCTTGA	3652
Qy	2693	CTGTGTTTCAAGATGAGGTGAAACAGCTACAGATTGACTACAAAAATCACGAATATCTGA	2752
Db	3653	CATGTTTCCGGGGGTGGTTAAGCAATTGCAAAATGCACTATCCCGGACATGAAAGTATGA	3712
Qy	2753	CTGGCGCTGCAATGCCAGAGCTTACGGGAAAGGCCGTTTATGCTGTCAAGGTACAAAGTCA	2812
Db	3713	CAGCCGGCGCTTCAAAAGGCTTAACCAAGAAAGAGTGTATGCCGTCGCGCAAAAGTCA	3772
Qy	2813	ACGAGAATCCACTCTACTCCGAGACTCTGAGCAGTGAACGTGTACTTACACGACGACAG	2872
Db	3773	ATGAAACCACTGTACGCAATCACAATCAGACATGTGAACGTGTGCTTCAACCCGCACTG	3832
Qy	2873	AAAAAGCAATTTGTGGAAGACGCTAGCTGTGTATCCCTGGATTAAGACACTTACAGCTA	2932

3833 AGGACAGCTAGTGTGAAAACTTTCAGAGGCGACCCATGATTAGACGCCACTAACA 3892
3893 AATATCCCGGGGATTTTCAGCGCTTCATTGACGATGCGACGCCGAAACAGAGCCATT 2992
3893 TACCTAAGGAAATTTTCAAGCTACTATAGAGCATGCGAGCTGAAACAGAGGAAATA 3952
2993 TGGCAGCGCTTCTTGAATAGCCGACAGCTGATGTGTTCAGAAATAGGTGAACGCTC 3052
3953 TTGTGTCAATTAACAGGCCCACTCCCGTGCATTCGCTTCAGTCGACAGCAACGTTT 4012
3053 GCTGGGGAAGGCTTTAGAGCGAGTCTTGGCAGCGCCAACTGTGTCTGACGACAGC 3112
4013 GCTGGGGAAGGCTTTAGAGCGAGTCTTGGCAGCGCCAACTGTGTCTGACGAGTCC 4072
3113 AGTGG---GAGAAGTTCACCATTTCAAGCATGACAGAGGCTACTGAAATGGCAC 3169
4073 AGTGGAGCAATGTTTCCACAGTTTGGGATGACAAACACATTCGGCCATTAAAGCTT 4132
3170 TGAACCTTCTTTGACAGGCTTTCTTGGAGTGAAGCTGACAGTGGGTTATTTCCGCTC 3229
4133 TAGACGTAATTTGATTAAGTTTTCGGCATGACCTTGACAGGGAAGTCTTTCTAAC 4192
3230 CTACCGTGCATTTACTTAACG-----GGATCAGCATGGGATA 3268
4193 AGAGCATCCCACTAACCTACCATCCCGCATTCAGCGAGGCCGGTACATTTGGGACA 4252
3269 ACTCGCAGGGAAGAATGATGTGGGCTTAATAGAGGTAGCAAGAGTGTACAGGC 3328
4253 ACAGCCGGAACCCGAGATGTGGTACGATCAGCGCATTTGCCGGAAGCTCTCCGTA 4312
3329 GATATCCGTGCATCAAAAACGGGTTGACACAGCAGGAGTACTGATATAGAAATATA 3388
4313 GATTTCCGGTGTCCAGCTAGC---TGGAAAGGCAACAATTGATTTGACAGCGGGA 4369
3389 CCATCAAGAGTACTCTTCCAAATTAATGTGTTCCATTAAATCCGCGGTTCCCACT 3448
4370 GAACCAAGTATCTCTGACAGCATTAACCTGTCCCGGTGAACCGAATCTTCTCAG 4429
3449 CGTTGATCGTTGACCAAAAAGGAGTACATGATCAGCGGATTCCTATCTAGA 3508
4430 CTTATGTCCTCGAGTACAGAGAGACACCCGCCCGGTCAAAAATTTCTTGAACAGT 4489
3509 TGAAGGCAATCTGTGTGTGTATC---GGCGATCTATCAGCATTCAGAGAGAAAG 3565
4490 TCAACACCACTCACTGATCTGTGTATCAGAGGAAAAATTTGAAGCTCCCGTAAGGAA 4549
3566 TAGAGTCCATGGGTCTATGCCCCAC-----TAATACATCAGGTGTGATCTGATTTGG 3619
4550 TCGAATGATCGCCCGATGCGATAGCCGGTGCAGATTAAGAACTAGAACTGCTTTCG 4609
3620 GAATACCTAGCANTGTGGTAAATATGATCATTTATCTTGTCAATGTTAGACCCCGTACA 3679
4610 GGTTCGCGCGGACGACG---GTACAGACTGGTGTCTATCAATGGAATTAATTAACA 4666
3680 GGAACATCACTACCAACAGTGCAGAGTCAAGCTATCCACACAGAGATGTTAACTGTA 3739
4667 GAAACACCACTTTCAAGAGTGCAGAACATGCGGAGCCTTAAAAACCTTTCCGGTT 4726
3740 AGGCTGTCCACCACTGAACATGCGGGAACATGTGTGATTAAGGTATGGGCTTCTG 3799
4727 CCGGCTCGAATTTGTTAACTCAAGAGGACCTCGGTGTGAAATCTCATAGGCTACCGC 4786
3800 ATGGCGCAACCGAATATATCATCTGCGGTGACGCTCATTTAGTTTACCGGTCT 3859
4787 ACCGCAACAGTGAAGAGTATGATCAGCTCTTTCAGAAAGTTTGTCAAGGTGTCTCAG 4846
3860 GTCAGCTTAAGAACATGCGGAAATATCTGAGTCTCTTCTGTGTTCTTCCGCAAGACA 3919
4847 CGAGACAGATTTGTGTCAAGCAATACAGAAATGTACCTGATTTTCCGCAACTAGACA 4906
3920 ACCGCAACCAACATGACAGGACAGACTCGGTGTATGCTTGAACAACATCTAT-CAA 3978
4907 ACAGC-CTGTACACGCAATTCACCCCGACCATCTGAATGCGTATTCGTCCGTGAT 4965

3979 GGGTCAACAGATACGAGCAGGAGAGACTCCAGCGTACAGAGTATCAGAGGTGACATT 4038
4966 GAGGTACAAAGATGAGATTTGAGACCGCCCGCTCATCCGACAAABAGGAAATATT 5025
4039 AGCAAGCGCTGACCAAGCTATCTTAAATGTCTTAATAGCAAAAGTCAACAGTTC 4098
5026 GCTACTGTCAAGAGAAAGATGTTCACGACGCCAATCCGCTGGGTAGACAGCGAA 5085
4099 GAGTGTGCGGTGACACTGTACCGAAAAATGGCCGCTCTTTGATAGACGCCAATGCT 4158
5086 GAGTGTGCGGTGACACTGTATTAACGTTGGCCACCAAGTTTACCGATTACGCCAG 5145
4159 GTCGGAGCGCTAGACT---TGTAAAGCAAGAACCGCTCATATACATGCTGTAGAGACC 4215
5146 ACAGGACCGCAAGAAATGACTGTGTGTCTAGAAAGAAATGATCCAGCGGTGCGCTT 5205
4216 AATTTTCTAAGATGCCGAACCGGAGGCGACCTTAAGCTGCGAGCTGTACATAGC 4275
5206 GATTTCCGAAGACCCAGAAAGCAAGAGCTTGAATTTGCTAACAAAGCCTACATGCA 5265
4276 ATAGCTTCATCTGCAACGCTGAGCGGATTAACAAAAATATCAGTACCGCTACTCAAC 4335
5266 GTGCGAGACTTAATGAATGAACATACATCAAGTCTGTCCCATTCACCTGCTATTA 5325
4336 GGCATTAATCTGTGTGCAAAAGATCGAGTATGCAATTCGATCCTGTTCACTGCT 4395
5326 GGCATTTAGCGACCGGAAABAGACCGCTTGAAGTATCACTTAATGCTTGTGACACCG 5385
4396 TTGCACACTACGAGTGCAGTGTCAACATATTTGCTTGTGAATAACATGAGAGACGAG 4455
5386 CTAGACAGACTACACGCGGAGTAAACCATCTATTTGCTGTGATTAAGAGTGAAGAGAA 5445
4456 ATATGAGGCGCATTCACCCGCAAGAAAGGTGGAATTT---GATGATGACAAAGCA 4512
5446 ATGACGCGGCACTTCAACTTAAGAGTCTGTAAACAGAGTGAAGATGATGAG 5505
4513 GTAGACATTAATGCTGTGTCAGGTCACACCAAGAGCTTTTGGCAGGACACAGTAAAC 4572
5506 ATGACGATAGTATGATGATGATTCATTCAGACAGTGTCTTGAAGGAAABAGAGATT 5565
4573 TCGTCATAGAGGCAAGTGTATTCATACCTGGAAGTACAGGATTCATCAGACCGCC 4632
5566 AGTACTCAAAAGAAATTTGATTTGCTGTTGGAAGGACCAAAATTCATTAAGAGCA 5625
4633 AAGGACTTCCGAAATTCATGCAATGTGCGCCCAAAATCTGAGGCTAATGACCAATT 4682
5626 AAGGACATGCGGAGATTAAGATGCTGTCTTCAATGACAGGAAAGTATGAACAATG 5685
4693 TGCTTGACATCCCTGGGAGAGATATGTCCAGATCCGCTCCAAATCCAGTAGAGAG 4752
5686 TGTGCTACATATTTGGTGAACCATGAGAGCAATCCGCAAAAGTCCCGGTGACCAT 5745
4753 TCAAGGCGCTGTCTCCACTCAACACTTCATGCTGTGTAATTAAGCTATGACGCT 4812
5746 AACCGGTCTAGCGCCGCCCAAAAGCTTGTGCAATGATGATGACATGACGCA 5805
4813 GAGCGGTATACAGTGTGCGCTCTGCGAAGAAAGACAGTTCCCGGTATGCTCATCTTC 4872
5806 GAAAGGCTCACAGACTTAAGCAATTAACGTCAAAAGAAATTCACATGATCTCTCCACC 5865
4873 CTGTGCGCAAGTACAGGATCACAGGCGTGCAGAGTACAGTGCAGCAACAGTCTG 4932
5866 CCCCTTCTTAAGCAAAATTAAGATGTTCAAGAGTTTGAAGTCAAGAAAGTATGCTG 5925
4933 TTTTCAAGCGCTGTACCAACCGGCTGTACACCCAGAGAGTACCGGAAAT 4982
5926 TTTAATCCGACACTCCCGCATTCGTTCCGCGCTTAAGTATCAATAGAGT 5975

RESULT 14
US-09-275-883-1
; Sequence 1, Application US/09275883

Publication No. US20030053988A1
GENERAL INFORMATION:
APPLICANT: Renner, Wolfgang A.
APPLICANT: Nieba, Lars
APPLICANT: Boersma, Marco
TITLE OF INVENTION: Inducible Alphaviral Gene Expression System
FILE REFERENCE: 1700.0020001
CURRENT APPLICATION NUMBER: US/09/275,883
CURRENT FILING DATE: 1999-03-25
EARLIER APPLICATION NUMBER: US 60/079,562
EARLIER FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 1
LENGTH: 11282
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: cDNA
US-09-275-883-1

Query Match 25 %; Score 1295.6; DB 11; Length 11282;
Best Local Similarity 56.0%; Pred. No. 0;
Matches 2816; Conservative 0; Mismatches 2134; Indels 80; Gaps 16;

QY 23 ATATGAGAAGATTACGCTTGATGATGCTGACAGCCGATGCTCAAGTGGTAAAGC 82
DB 956 AGAAGCCAGTGTAAACCTAGACGTAGACCCCAAGTCCGTTTCTGTCACACTGCAAA 1015
QY 83 GAGCGCTTCCCAATTGATGATGAGAGGAGGAGGCTGCTGACATGACATGACATG 142
DB 1016 AAGCGTCCCGCAATTTAGAGTAGACAGCAGAGTCACTCAATGACATGCTAATG 1075
QY 143 CCAAGCGCTTTTCCGATGCTGGAACAAGCTCATTTGAGAGGAAATCCAGCGGACCAAG 202
DB 1076 CCAAGCGCTTTTCCGATGCTGGAACAAGCTCATTTGAGAGGAAATCCAGCGGACCAAG 1135
QY 203 TTATCTTGACATTTGGAAGTGGCGCCGTCAGCATGACATTCATCAATCCGCTATCAT 262
DB 1136 CGATTTTGGACATGAGGACCGGACCGGCTCTGTAAGATGTTTTCAGACACCATATCAT 1195
QY 263 GTATCTGCTTATGATGATGAGGCTGAGACCCGAGACATCAACAGGATGAGAGAAC 322
DB 1196 GTGCTGCGCCATGCTGATGTCAGAGAACCCGAGACCGATGATGAATACGCACTAAC 1255
QY 323 TTAGAAAA-----GTGACATTACCGACAGAACATAGCTCTTAAAGCGGACGCTGC 376
DB 1256 TGGCGGAAAAAGCGTGAAGATTACAAACAAGAACTTCATGAGAAAGATTAAAGATCTCC 1315
QY 377 TGGAGTCATGTCAACACAGACGAGACGCTCATCTCTGTGATGACACAGACGCA 436
DB 1316 GAGCGCTACTGATAGCGCGGATGCTGAACACCATCGCTTCTTTCACACAGATGTTA 1375
QY 437 CGTGTAGTACTTTGGAAGTGAAGTGAATCAAGATGTGTACGACAGTCCATGACCGCA 496
DB 1376 CCGTGAACATGCGTCCGAATATTCGTCATGCAAGAGGTGA---TATCAACGCTCCG 1432
QY 497 CATCAATTAACCAAGGCGCTTAAAGAGTGAAGCAATTTACTGATAGGCTTTGCA 556
DB 1433 GAACATATCATCATGAGCTATGAAGAGCGTCCGACCTGTACTGAGATGCTTCGACA 1492
QY 557 CGACCCCTTTATGTAACAAAAATGAGGCTTCTACCTACTACAAACGAACTGGG 616
DB 1493 CCACCCAGTATGTTCTCGGCTATGCGAGGTTCTATCCCTGCGTAAACACCACTGG 1552
QY 617 CTGACGAGAGATATGGAAGACGCTAACTTTGCGCTGCTGATCACTTCAGAGCA 676
DB 1553 CCGAGAGAAAGCTCTTGAAGGCGTAACTGCACTTTGACAGCAACAAGCTGAGTAG 1612
QY 677 GAGGCTTGAAGAACTCTCAATCTTGAAGAGAAAGGCTCCAACTCTAATTAAGATCA 736
DB 1613 GTAGAGCAGAAATTTGTGCAATTAATGAGAGAAAGAGTGAAGCTCCGCGTCCGCGTTT 1672

QY 737 TATCTCGGTTGGTTCAACAACTACACAGAAATGATGATCACTGTACGATGCGGATC 796
DB 1673 ATTCTCGGTAGATGACACATTTATCCAGAAACACAGCCAGCTTGGACAGCTGGATC 1732
QY 797 TTCCAACGTGTTCCTCACTTGAAGAAAGTCTAACTTACAGATGATGAGGACCAATG 856
DB 1733 TTCCATCGGTGTTCCACTGATGAAAGACAGTGTGATCACTTGGCCGTGTATACATG 1792
QY 857 TCAGCTGAGAGGATGCTCATCAAAAGATGATGATGATGATGATGATGATGATGATG 916
DB 1793 TGAGTTGGAAGGCTACGATGATGAAAGAAATCACTATGATCCGCGATCACGGAAGAA 1852
QY 917 TTGAGAACTTGGCGCTCAACATGATCGGAGGTTTCTTGAAGTTCAGAAAGTCAAGATA 976
DB 1853 CCGTGGATTAAGCGGTTTACACAAATGCGAGGCTTCTTGTATGCAAAAGTATGACA 1912
QY 977 CGCTGCGCGGAGAGGTTTCTTGTGTGTGATGATGATGATGATGATGATGATGATG 1036
DB 1913 CAGTAAAGAGAAACGGGTATGCTTCCCTGTGTGACGTATCCCGGACCATATGCG 1972
QY 1037 ATCAGATGACAGGATTTGCGCACTGACCTGATGATGATGATGATGATGATGATGATG 1096
DB 1973 ATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2032
QY 1097 TTGGGCTCAACCAAGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1156
DB 2033 TTGGGCTCAACCAAGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2092
QY 1157 AGAATATCTATTAACAGTGTGCTCCGAGCGTTTTCAGAGTGGCGGCTGAATCTGTG 1216
DB 2093 AAAATTAATCTTCTGCGCATATGACACAGAGTTTACAGAAATGGCTTAAGAGCCAGAG 2152
QY 1217 CCGACTTGAAGAGAGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276
DB 2153 ATGATCTTGTATGAGAGAAATGCTGGTACTAGAGAGAGAGAGATGATGATGATGATG 2212
QY 1277 GCTGGGCTTTCAGAGCCAGAAATACATCATCTTCAAGAGGCTGATGAGAGAGAGAG 1336
DB 2213 TGTGGGCTTTCAGAGCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 2272
QY 1337 TTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1396
DB 2273 GGTAAAGTCCAGAGCTCTTTAGCGCTTTTCCATGCTGCTGATGAGAGAGAGAGAGAG 2332
QY 1397 TCGATGAGGCTTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1453
DB 2333 TGCCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2392
QY 1454 CACGCGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1513
DB 2393 TGCTGAGTCTCGAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 2452
QY 1514 AAGTGTGACAGCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1573
DB 2453 AGAAGCGAGAGCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2512
QY 1574 AAGAGACCTGAGAGCA-----GAAGTATGATGATGATGATGATGATGATGATGATG 1624
DB 2513 TCGAGGAGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2572
QY 1625 GTAGGATGAGAGACACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1684
DB 2573 CATTAATGTAATCCCGCGGCTACGTAAGATTAATCTTCAAGCAATTAATGATGATGATG 2632
QY 1685 TTGGCTTACGATATCTTCAACCGAGGCTTATGATGATGATGATGATGATGATGATGATG 1744
DB 2633 TCGGAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2692
QY 1745 TCGACCATTTGGCGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1804
DB 2693 CGACCGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2752
QY 1805 AAGTCAGACATACACAGGTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1864

```
Db 2753 CGGTCGAACCATGACGAGCTAAAGTACTGATGCCAGGAGGCGCGTACCATGGCCAG 2812
Qy 1865 ACTTCCAGGCGATTGAGTGAAGGCGCTACGATCGTTTTCACAGAGGAGATTCGTAACA 1924
Db 2813 AATTCCCTAGCACTGAGTGAAGGCGCACGTTAGTGTACACAGAAAGAGATTTGTGAACC 2872
Qy 1925 GATACCTGCACCATCGCATCAACGAGGAGCGCTAAACACTGACGAAAGTACTTA 1984
Db 2873 GCAACCTATACCACTATTCGATGATGCGCCCGCCAGAAATACAGAAAGAGAGCAATCA 2932
Qy 1985 AGACTGTAAAGACTCAGAGACAGACTCAGAAATACGTTCTGATATTGACGACGAAAGT 2044
Db 2933 AGGTTACAAAGGACAGAGCTTCAAGAAACAGAGTACGTTTGAAGTGAACAAGAGCGTT 2992
Qy 2045 GTGTTAAGCCAGAGACGAGAGTCCCTTGTGCTTACCGGTGATCTGTATATCAACAT 2104
Db 2993 GCGTTAAGAAAGAAAGAGCTCAGAGTCTGTCTCTCGGAGAACTGACCAACCTCCCT 3052
Qy 2105 TTCAGAGTTTGCCTACGAGAGTCTCAAGACAGCAGCAGCAGCCTCAAAAGTCCCA 2164
Db 3053 ATCTAGAGCTAGCTCTGAGGAGATGAAAGACCACTGCGGTCCCTACAGAGTCCAAA 3112
Qy 2165 CCATCGAGTCTATGAGAGTCCAGGTTCAAGTAAATCTGAATCATCAAAAGCGCTGTA 2224
Db 3113 CAATAGAGGTATAGGACACCGGGGTCCGCAAGTCAAGTATTATCAAGTCACTGTCA 3172
Qy 2225 CTAAGAAAGATCTGCTGTGAGTCCGAAAGAAAGAAATCGCGCAAGAAATCATAGGATG 2284
Db 3173 CGGACAGAGATCTGTGTTACCGCGGAAAGAAAGAAATTTGCGGAAATTGAGGCCACG 3232
Qy 2285 TAAAGAGATGAGAGATGATGATGTGCTAGAGACTGTGATTCAGTGTCTTCTAAATG 2344
Db 3233 TGCTAAAGACTTGAAGGAGTATGAGATTAAGTGAAGACAGTATTCGTTATGCTCAAG 3292
Qy 2345 GGGTTAAGACCCCGTTAAACCTGTGATCATGTAGAGCATTTGCTGCCATGCAAGGA 2404
Db 3293 GATGCCCAAAAGCCGTGAAGAGTGTGATGTTGACGAAGCGTTGCGCGTCCACAGAGAG 3352
Qy 2405 CGCTGCTGCGACTGATATTCCTGCTCAACCC---TAAAGAAAGTGTATTTGCGGGAGC 2461
Db 3353 CACTACTGTGCTTATGCTATGCTATGCTCAGGCCCCGCAAGAAAGTATGATGCGGAGC 3412
Qy 2462 CAAACATGCGGCGCTTCTTAAACATGATGCTGAAAGTACTATTTAAAC----- 2512
Db 3413 CCATGCAATGCGGATCTTCAACATGATGCACTTAAAGTACATTTCAATACCTGAAA 3472
Qy 2513 ATGACATATGCACTGAAGTGTACATTAAGCATCTTAGAGGTGACACAGACTGTAA 2572
Db 3473 AAGACATATGCAACAGACATTTCTAAAGTATATCTCCGCGTTCACACAGCCAGTTA 3532
Qy 2573 CCGGCAATGCTCCACGCTCTTCTACGACAGCGAATGAACGTTAAACCATGTGCTG 2632
Db 3533 CAGCTATTGTATCCAGACTGCACTGATTAAGTGAAGATGAACCAAGAACCCGTGCAAG 3592
Qy 2633 ATAAATCATCATATGATACACAGGAGCAACAAGCGCACAAGATGATCTGATTTCAA 2692
Db 3593 AGAACATTTGAATGATATTTACAGGGGCGCAAAAGCCGAAGGGATATCATCTGTA 3652
Qy 2693 CTTGTTCAAGAGATGGGTGAAGAGCTACAGATTTGACTCAAAAATCAAGAAATCATGA 2752
Db 3653 CATGTTCCGGGGTGGGTTAAGCAATTCGAAATTCGCGAGACATGAAGTAATGA 3712
Qy 2753 CTGCGGCTGCACTCCAGAGACTTACGCGGAAAGCGTTATGCTGTCAAGTACAAAGTCA 2812
Db 3713 CAGCGCGGCTTCAAGAGGCTTAAACAGAAAGAGGTATGCGTCCGCGCAAAAGTCA 3772
Qy 2813 ACGAATTCACATCTGACGAGACTTCTGAGCACTGGAAGTGTACTTACAGGACAG 2872
Db 3773 ATGAAACCCACCTGTACCGGATCAATGAGCAATGTGAAGTGTGCTTCCACCGGACATG 3832
Qy 2873 AAAAAGCATTTGTGGAAGAGCGTAGCTGTGATCCCTGATTAAGCACTTACAGCTA 2932
Db 3833 AGGACAGGCTAGTGTGAAAACTTTCAGGGGCAACCCATGATTTAGCACCCCACTACA 3892
Qy 2933 AATTCCCGGGGANTTTCAGCGCTTCAATTGAGACACTGGAGCGGGAACAGAGCCATTA 2992
Db 3893 TACTTAAAGAACTTTCAGGCTATATAGAGACTGGGAAGCTGAACACAGGGAAATTA 3952
Qy 2993 TGGCACCGCTTCTGATTAACCGGACAGACTGATGTGTCCAGAAATTAAGTGAACGCTC 3052
Db 3953 TTGCTGAATTAACAGCCCACTCCCGTCCATTCGTTTCAGGTGCAAGCCAAAGCTTT 4012
Qy 3053 GCTGGGGAAGGCTTTAGAGCACTTGGCCACAGCGCAACATTTGTCTGACGAGACAGC 3112
Db 4013 GCTGGGGAAGCAATTTGGAACCGATCTACACAGCGCGGTATGTTACTTACCGGTTGCC 4072
Qy 3113 AGTGG---GAGAGCTTCCACCATTCGAAGATGACAGAGGCTACTGAAATGGCAC 3169
Db 4073 AGTGAAGCAACTGTTTCCACAGTTTTCGAGTGAACCAACATTCGCGCATTTAGCCT 4132
Qy 3170 TGAACCTCTTTTGACACCGGTTCTTTGAGTGAAGCTGACAGTGGTTATTTCCGCTC 3229
Db 4133 TAGACGTAATTTGATTAAGTTTTTCGCAATGACCTTGACAGGGAATGTTTCTAAAC 4192
Qy 3230 CTACCGTGCATTTACTTACAG-----GGATCAGCACTGGGATA 3288
Db 4193 AGAGCATCCCACTAACCTATCCGCGCATTCAGCAGGCGCGTGTGCTATTTGGGACA 4252
Qy 3269 ACTGCGAGGGAAGAAATGATGATGGGCTTAATAGAGGTAGGCAAGAGATGTACAGGC 3328
Db 4253 ACAGCCAGGAACCCGCAATGTAGTACAGATCACGCAATTCGCGGAACTCTCCGTA 4312
Qy 3329 GATATCCGTCATCACAAAAGCGGTTGACACAGCAGGAGTATGATATAAGAAATATA 3388
Db 4313 GATTTCCGGTGTTCACATGAC---TGGAAAGGCAACAACCTGATTTGACAGCGGGA 4369
Qy 3389 CCATCAAGACTACTCTTCAACATTAATGTGTTCCATTAATCCCGGTTGCCCACT 3448
Db 4370 GAACCAAGTATCTCTGCAAGCATTAACCTGTCCCGTGAACCGCAATCTTCTCAG 4429
Qy 3449 CGTTGATGCTTGAACCAAGAGGACAGGTTACACTGATCAAGCGGATTCCTATCTAGA 3508
Db 4430 CTTAGTCCCGAGTACAGAGAGACACCCGCGCTGCAAAAATTTGTAACAGT 4489
Qy 3509 TGAAGGCAATCTGTGTTGTGATC---GGCATCTTATCAGATTCAGGGAAGAAAG 3565
Db 4490 TCAACACACACTAGTACTTGTGTATCAGAGAAATAATGAAGTCCCGTAAAGAA 4549
Qy 3566 TAGAGTCAATGGTCTCTTGCAC-----TAATACATCAGGTGATCTGATTTGG 3619
Db 4550 TCGAATGATTCGCCCCGATTTGAGATAGCCGATGCAATTAAGAACTCAACCTGCTTTGG 4609
Qy 3620 GAATACCTAGCATGTGCGTAAATATGACATTTACTTGTCAATGTAGGACCCCGTACA 3679
Db 4610 GGTTCGCGCGCAGGACG---GTACGACTGTGTTTCACTACATTTGAACTTAATACA 4666
Qy 3680 GGAACCATCACTAACCAACAGTGCAGAGATCAGCTATCCACACAGCATCTAACGTGTA 3739
Db 4667 GAACACACACTTTCAACAGTGGAAAGACCATGCGGACCTTAAAAACCTTTGCGGTT 4726
Qy 3740 AGGCTGTCCACACTGTAACACTGGCGAAACATGTGTGCTATAGGATATGGCTTCTG 3789
Db 4727 CCGGCTCTGAATTTGTTTAACTCAGAGGCAACCTCTGCTGAAGTCTTATGGCTACCCG 4786
Qy 3800 ATGCGGCAACCGGAATATCATCATCAGCGGTGGACGCTCAATTAGGTTAACCGGTCT 3859
Db 4787 ACCGCAACATGTAGAGAGTGTACCGCTTTCGCAAAAGTTGTACAGGTGTCTCAG 4846
Qy 3860 GTCACTTAAGAACACTGCCGAATATCTAGAGTCTCTTGTGTTCTTCCGCAAGACA 3919
Db 4847 CGAGACAGATTTGTGTCTCAAGCAATACAGAAATGTACGATTTTCCAGCACTAACA 4966
Qy 3920 ACGGCAACCAACACATGACAGAGCAAGACTCGGTGATAGTCTTGACACATCTAT-CAA 3978
Db 4907 ACAGC-CGTACAGGCAATTCACCCGACCACTGTGAATTTGCGATTTGCTCGTGTAT 4965
```

OY	3979	GGGCTCAACGAGGTACGAGGAGGGGAGAGCTCCAGGCTGACAGAGTATCAGAGGTGACATT	4038
Db	4966	GAGGGTACAAAGATGAGAGTTTGAGCCGCCGCTCATACCGCACCAAAAGGGAAGATATT	5025
OY	4039	AGCAAGACGGCTGACCAAGCTATCGTTAATCTGCTAAATPACAAAGTCAACAGTTTCC	4098
Db	5026	GCTGACTGTCAAGAGGAAGCAGTTGTGTCAACGACGCAATTCGGCTGGGTAGACAGGGGAA	5085
OY	4099	GGAGTGTGCGGTGCACTGTACCGGAAAATGCGCGGTGCTTTTGATAGACGCCAATAGCT	4158
Db	5086	GGAGCTGCGCGTGCCTATCTATTAACAGTTGGCGCAACAGTTTACCGATTCCAGCCACGAG	5145
OY	4159	GTCGGGACGGCTAGACT--TGTGAAGACGAAACCGGCTCATCATAGTCTGTAGAGCCC	4215
Db	5146	ACAGGCAACCGCAAGATATACGTCTGTGCTTAGAAAAGAAAGTATCCAGCGGTGGCCCT	5205
OY	4216	AATTTTCTTAAGATGCGCGAAACCGAGGGCGACTTAAGCTCGCAGCTGCTTACATGAC	4275
Db	5206	GATTTCCGGAAGACCCGAGAGCAGAACTTTGAAATTGTTACAAAAGCCCTTACATGCA	5265
OY	4276	ATTAGGTCCTCATCGTCAACGCTGAGGGGATTACAAAATATCAGTACCGCTACTGTCAAC	4335
Db	5266	GTTGGAGACTTAAGTAATGAACTATPAACATCAAGTCTGTGCGCATTTCACTGCTATCTACA	5325
OY	4336	GGCATCTTATTTCTGGGCGCAAGATGCAATGCAATGTCATTCACCTGTTACATGCT	4395
Db	5326	GGCATTTTACGAGCCGGAAGAAACCGCTTTGAAGTATCATCTTAACGTGTTACAAACCGCG	5385
OY	4396	TTTGCACACTACGAGATGCCGATGTCAACATATATGCTTGGATPAAACAATGGAGACGAG	4455
Db	5386	CTAAGACAGAACTGACGCGGAGGTAACTATTCATTTGCTGGATTAAGATGGAAAGAAAGA	5445
OY	4456	ATAATTCAGGCGCATTCACCGCAAGAAAGAAAGCTCCAAATTC---GATGATGACAAGCCA	4512
Db	5446	ATCGACGGGCACTTCCAACTTAAGAGTCTGTAAACAGCTGAAAGATGAAGATATGAG	5505
OY	4513	GTAGACATTTGACTTGTGTCAGGGGTCAACCCAAACAGCTTTTGGCAGGCGACACAGTTAC	4572
Db	5506	ATCGACGATGAGTTAGTATGGAATTCATCAACAGATTTGCTTGAAGGGAAGAAAGGATTC	5565
OY	4573	TCCGTCATATGAGGGCAGTTGTATTATCATCTCGAAGGTACACGATTCATCAGACCGCC	4632
Db	5566	AGTACTCAAAAGGAAATTTGATTATTCGTACTTCAGAGGCACCAATTCATCAAGCAGCA	5625
OY	4633	AAGGACATTTGCCGAATCATGCAATGAGGCCCAACAAATCTGAGGCTAATGACAGATT	4692
Db	5626	AAAGACATGCGGAGATTAAGGTCTGTTCCTTATGACAGGAAAGTAAAGAACCACTG	5685
OY	4693	TGCTTTGATCATCTCGGGGAGAGTATGTCCAGATTCGCGTCCAATGCGCCAGTAGAGAG	4752
Db	5686	TGTGCGTATCATATTGGGTGAGAACCATGGAACCAATCCGCGAAAGTCCCGGTGACCAT	5745
OY	4753	TCAAGGCGGTGTGCTCCACTCAACACTTCATGCTGTGTATTAATACGTATAGGACT	4812
Db	5746	AACCGCTGTCTACCGCCGCAAAACGTGGCCGTGCTTTGCAATGATATGCATATCCCA	5805
OY	4813	GAGCGCGATACAGATTGCGCTTGCGGAAGAAAGACGTTCCGCGATATGCTCATATTC	4872
Db	5806	GAAAGGCTCCACAAGCTTAAGAACATATACGTCAAGAAAGTTACAGTATGCTCTTCAAC	5865
OY	4873	CTGTTGCGGAAGTACAGGATCACAGCGCTGACAGAGCTACAGTCAGCAAAACAGTCTG	4932
Db	5866	CCCCCTCTTAAGCAAAAATTAAAGATTTTCAAGAGGTTCAGTCAAGAAAGTATGTCCTG	5925
OY	4933	TTTTTCAGGCGTGTATCCACCGGCTGTACACCCCAAGAAATGTCGGGAAAT	4992
Db	5926	TTTAATCCGACACTCCCGCATTTGCTTCCGCGCGTAAAGTACATAGAAAGT	5975

RESULT 15
US-10-123-101-51
; Sequence 51, Application US/10123101

```

; Publication No. US20030148262A1
; GENERAL INFORMATION:
; APPLICANT: PERRI, John
; APPLICANT: PERRI, Silvia
; APPLICANT: THUDUM, Kent
; TITLE OF INVENTION: CHIMERIC ALPHAVIRUS REPLICON PARTICLES
; FILE REFERENCE: 2300-17924 / PP17924.002
; CURRENT APPLICATION NUMBER: US/10/123,101
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 60/295,451
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: VEE nsp fragment oligo 1
; US-10-123-101-51

```

Query Match	6.6%	Score 330.8;	DB 12;	Length 705;
Best Local Similarity	67.0%;	Pred. No. 1e-97;		
Matches 470;	Conservative 0;	Mismatches 232;	Indels 0;	Gaps 0

QY	627	AGATTGGAAACAGTAACTTGGCCCTGGTAACCTCAATCTTCAGAGAGACGGCTTGG	686
Db	3	AGAGTTAAACGGCTGCTAACAATAGCCCTTATGCAGCTTGAACGTTATGAGACGGTCAAGTGA	62
QY	687	AAAACCTCAATCCTTAGAAGAAAGAGGCTCCAACTTACTAATTAAGATCATATTCTCGT	746
Db	63	AGGAGATGCCATTTCTTAGAAAGAAAGTATTTGAAACCATCCAAACAAATGTTCTATCTCTGT	122
QY	747	TGGTTCAACATCTACAACAGAAATGATATCACTGTAGTGGCTGGCACTCTCCAAAGT	806
Db	123	TGCTCGAACCATCTTACACGAGAAAGAGGAACTTACTATGAGAGCTGGACCTGCCCTGT	182
QY	807	GTTCCACTTGAAGAAAGATCTAACTTCACAGTATGATGTGGACCAATGTTCAGCTGTGA	866
Db	183	ATTTCACCTTAGTGGCAAGCAAAATTTACATATGTCGGGTGTAGATTAATATGTTAGTGGCA	242
QY	867	AGGGTAGCGTATCAAAAAGATTAAGATCAGCCACAGAACTATACGTTAAGTTGAAACTT	926
Db	243	CGGGTAGCGCTTAAAGAAATAGCTATATAGTCCAGGCTGTATGGAAAGCCTTCAGGCTA	302
QY	927	GGCGTCCACATATGATCGCGAGGGTTCTTAGTTGCAAAGTTCACAGATACGCTCGCGG	986
Db	303	TGCTGTACATGATGACCGCGAGGGATTTCTGTGCTGCAAAATGACAGACACTTGAACGG	362
QY	987	CGAAGGGTTTCTTTGCTGTGTGACGTATGTACACAGCCACACTTTCGATCAGATGAC	1046
Db	363	GGAAAGGGTCTCTTTTCCCGTGCACAGTATGTGCACACTCATTTGTGTGACCAAAATGAC	422
QY	1047	AGGATTCGTGGCACTGACGTTAGTGTGGATGAGACGCAAAAACTAATTTGGTTGGGCTCA	1106
Db	423	TGCGATATCTGGCAACAGTGTCACTGATGCGGAACGACGCGAAAACTGCTGTGTGGCTCA	482
QY	1107	CCAAAAGATTTGTCTCAATGTGTAGACGCGAAAGAAATTAACAACAATGACAAATATCT	1166
Db	483	CCACCGTATAGTGTCTCAACGGTCCACCCAAAGAAACCAATATCCATGAATAAATTAAGT	542
QY	1167	ATTACAGTGTGCCCGCAGCGTTTTCAGGTGGCGCGGTAAACATCTGCGCACTTGGGA	1222
Db	543	TTTTCGCCGTATGGCCCAAGGCAATTTGTGAGGTGGGCAAAAGAAATTAAGAAATCAAGA	602
QY	1227	CGACGAGAAAGAACTAGAGGGGTGCGGAGCGCACTTTACTATAGGCTGTGCTGGGCTTT	1288
Db	603	AGATGAAGGCGCACTAGAGACTACAGAGATAGACAATGTATCATGTGGGCTGTGTGGGCTTT	662
QY	1287	CAAGATCCGAAAAATCAATCATCTACAAAGACCTGTATC	1328
Db	663	TAGAAAGCACAGATTAATCTTATTTAATACGCGCTCGGATAC	704

Mon Nov 17 10:39:22 2003

us-10-023-649-1_copy_1_5000.rnpb

Page 38

Search completed: November 16, 2003, 00:28:09
Job time : 1549 secs

Query Match 89.8%; Score 4492; DB 3; Length 11492;
Best Local Similarity 95.8%; Pred. No. 0; Mismatches 127; Indels 82; Gaps 18;
Matches 4826; Conservative 0;

QY 1 ACCCTCAAACTAATCGATCCAAATATGAGAAAGAAATTCACGTTGACTTATGATGCTGAGAC 60
DB 25 ACCCTCAAACTAATCGATCCAAATATGAGAAAGAAATTCACGTTGACTTATGATGCTGAGAC 84
QY 61 CCGTATGTCAAAGTGTTCAGCGGACGTTTCACAAATTTGAGATCGAAAGCAAGCGAGTTC 120
DB 85 CCAATATGTCAAGTGTTCAGCGGACGTTTCACAAATTTGAGATCGAAAGCAAGCGAGTTC 144
QY 121 ACTGACAAATGACCAATGCGCAATGCGAGCGTTTCGATGAGGAGCAAAAGCTGATGAG 180
DB 145 ACTGACAAATGACCAATGCGCAATGCGAGCGTTTCGATGAGGAGCAAAAGCTGATGAG 204
QY 181 AGCGAAGTCGACCGGAGCAAGTTATCTTGACATTTGACATTTGAGCGCCGTCAGACATGCA 240
DB 205 AGCGAAGTCGACCGGAGCAAGTTATCTTGACATTTGAGCGCCGTCAGACATGCA 264
QY 241 CATTCCAAATCACCGCTATCATTTGATCTGCCCTATGATAGCGCTGAAAGCCCGAGACA 300
DB 265 CATTCCAAATCACCGCTATCATTTGATCTGCCCTATGATAGCGCTGAAAGCCCGAGACA 324
QY 301 CTACAAGGTATGAGAAAGCTTAAGAAAGTGAACAT-TACGACAAAGACATAGGCTC 359
DB 325 CTACAAGGTATGAGAAAGCTTAAGAAAGTGAACATGATACCGCAAAATATAGGCTC 384
QY 360 T-AAGGCGGAGACCTGCTGAGAGTCAATGTCACACGAGCGAGAGATCCATCTCTGT 418
DB 385 TNAAGGGGAGACCTGCTGAGAGTCAATGTCACACGAGCGAGAGATCCATCTCTGT 444
QY 419 GTATGACACAGACGCCAGCTGTAGTACTTTGAGAGTGA-TGCAATATACCAAGATGTG 477
DB 445 GTATGACACAGACGCCAGCTGTAGTACTTTGAGAGTGAAGTATACCAAGATGTG 504
QY 478 TAGCGAGTCATGACCGGACATCAATCAACACGAGCGCTTAAAGAGTTAGACAATT 537
DB 505 TAGCGAGTCATGACCGGACATCAATCAACACGAGCGCTTAAAGAGTTAGACAATT 564
QY 538 TACTGATAGGCTTTGACACGCCCTTTTATGTATCAAAAACATGCGAGTTCTTACCCT 597
DB 565 TACTGATAGGCTTTTGAACAGACCCCTTTTATGTATCAAAAACATGCGAGTTCTTACCCT 624
QY 598 ACTTACAAACGAATGCGCTGAGAGAGATTTGGAAGCAAGTATGCGCTCGGT 657
DB 625 ACTTACAAACGAATGCGCTGAGAGAGATTTGGAAGCAAGTATGCGCTCGGT 684
QY 658 AACTCAGATCTTCAGAGAGAGGCTTGAAGAACTCAATCTCTTGAAGAGAGGCTC 717
DB 685 AACTCAGATCTTCAGAGAGAGGCTTGAAGAACTTCAATCTCTTGAAGAGAGGCTC 744
QY 718 CAACCTACTAATAAGATCATATTCGCTGTTTCAACATCTACACAGAAATATGATCA 777
DB 745 CAACCTACTAATAAGATCATATTCGCTGTTTCAACATCTACACAGAAATATGATCA 804
QY 778 CTGTTTACGTAGCTGATCTTCCAAAGCTGTTTCACTTGAAGAGAAAGTCTTATCTTCA 837
DB 805 CTGTTTACGTAGCTGATCTTCCAAAGCTGTTTCACTTGAAGAGAAAGTCTTATCTTCA 864
QY 838 GGTGAGATGAGGACCAATGTCAGCTGTAAGAGGTTCGATCAATCAAAAATAACGATCAGC 897
DB 865 GGTGAGATGAGGACCAATGTCAGCTGTAAGAGGTTCGATCAATCAAAAATAACGATCAGC 924
QY 898 CCAGGACTATATCGGTAAAGTTGAGAACTTTGGCGTCCAAATGATCGCGAGGGTTTCTTG 957
DB 925 CCAGGACTATATCGGTAAAGTTGAGAACTTTGGCGTCCAAATGATCGCGAGGGTTTCTTG 984
QY 958 AGTTGCAAAAGTCAAGATACGCTGCGCGGCGAGAGGGTTTCTTTGCTGTGTATCGTAT 1017
DB 985 AGTTGCAAAAGTCAAGATACGCTGCGCGGCGAGAGGGTTTCTTTGCTGTGTATCGTAT 1044

QY 1018 GTACAGACCACTATTCGATCAGATGACAGGAGTTTCGCAATGACGTTAGTGTGAT 1077
DB 1045 GTACAGACCACTATTCGATCAGATGACAGGAGTTTCGCAATGACGTTAGTGTGAT 1078
QY 1078 GACGCAAAAACCTATTTGTTGGGCTTCACCAAGATTTGTCTCAATGATGAGCGCAA 1137
DB 1079 -----GGGCTCAACCAACGATTTGTCTCAATGATGAGCGCAA 1117
QY 1138 AGAATATCTAACCAATGACGAATCTATTTTCCAGTGTGCTCCGAGGGGTTTCCAGG 1197
DB 1118 AGAATATCTAACCAATGACGAATCTATTTTCCAGTGTGCTCCGAGGGGTTTCCAGG 1177
QY 1198 TGGGCGCGTGAATCGTGTGCGCATTTGAGACGAGAGAAAGTATGAGGCTGCGGAGCGC 1257
DB 1178 TGGGCGCGTGAATCGTGTGCGCATTTGAGACGAGAGAGAGAGTATGAGGCTGCGGAGCGC 1237
QY 1258 ACTCTTACTATGAGGCTGTGCTGCGGCTTTCAAGACCCAGAAATACATCTCATCAAG 1317
DB 1238 ACTCTTACTATGAGGCTGTGCTGCGGCTTTCAAGACCCAGAAATACATCTCATCAAG 1297
QY 1318 AAGCTGTGTACGGAAC-AAATTAAGAAATGATCCTGCCGTC-TTGAATCTATTTGTGATTC 1375
DB 1298 AAGCTGTGTACGGAACAAATTAAGAAATGATCCTGCCGTCCTTGAATCTATTTGTGATTC 1357
QY 1376 C-ACGCTTACAGCCACGCGGCTCGAT-----ATGGGCTTCCGCGCT-AGGCTCAAGCT 1427
DB 1358 CAGAGCTTACAGCCACGCGGCGGCTCGAATATGAGGCTTTCGCGCTTNAAGGCTCAAGCT 1417
QY 1428 GCTGCTTGAACCAATCTGTCAAAACCCGACCGGCTATTAATATGCGGATGTGAGCATCT 1487
DB 1418 GCTGCTTGAACCAATCTGTCAAAACCCGACCGGCTATTAATATGAGGCTTGTGAGCATCT 1477
QY 1488 GCGTGGCTTACAGCAAAAGCTGAAGAGTGGCTGACG-CGAAAGATGACAGAAAGCCC 1546
DB 1478 GCGTGGCTTACAGCAAAAGCTGAAGAGTGGCTGACGCGGAGAAATGACAGAAAGCCC 1537
QY 1547 TGCCACCTTGCTCTCCCTGAATAGAAAAGAGACCGTAGAGGAGAGAGTATGACTCATTA 1606
DB 1538 TGCCACCTTGCTCTCCCTGAATAGAAAAGAGACCGTAGAGGAGAGTATGACTCATTA 1597
QY 1607 TGCAAGAGGAGAGAGAGTATGCTGAGACACC-ACGAGACATCAGGCTGACAAAT 1665
DB 1598 TGCAAGAGGAGAGAGAGTATGCTGAGACACCAACGAGAGACATCAAGTATGACAAAT 1657
QY 1666 TACCCAGGCG-AAAGAGAGATGGGCTTGA-CGCTATCTTTTCAACCCAGGCGGATTT-G 1722
DB 1658 TACCCAGGCTGAAAGAGAGATGGGCTTGA-CGCTATCTTTTCAACCCAGGCGGATTTAN 1717
QY 1723 AATAGTAAAAAACTGGGCTATTCACCCATTTGGCGGAGCAAGTACTGTATATGACTC 1782
DB 1718 AATAGTAAAAAACTGGGCTATTCACCCATTTGGCGGAGCAAGTACTGTATATGACTC 1777
QY 1783 AAAGTAGGGGAGAGAGATACAAAGTGTGAGCCATACCAAGGTAAAGTATTTGA-TCCAGA 1841
DB 1778 AAAGTAGGGGAGAGAGATACAAAGTGTGAGCCATACCAAGGTAAAGTATTTGA-TCCAGA 1837
QY 1842 AGGAGCGGC-GGTCCCTGTTCAAGACTTCCAGGACTTGAAGTGAAGGCGTACGATCGTTT 1900
DB 1838 AGGAGCGGCAGGCTCTGTTTCAAGACTTTCAGGACTGAAGTGAAGGCGCAAGTCTGTTT 1887
QY 1901 TCAACGAGAGGAGTGTGTAA-----CAGATACCTGCAACCAATCGCAATCA 1948
DB 1898 TCAACGAGAGGAGTGTGTAAACCAAGATTTTTCACCCATGCAAGCTTTCAACCTA 1957
QY 1949 ACGGAGAGCGCTAACAATGACAGAGATCTTAAGACTGTAAAGCTTCAAGCAACAG 2008
DB 1958 TAGTGAATCGCTATTAATGACAGAGATCTTAAGACTGTAAAGCTTCAAGCAACAG 2017
QY 2009 ACTCAGAAATCGTCTTGTGAATTTGACGCAAGAAAGTGTAAAGCAAGAGAGAGCAGGTC 2068
DB 2018 ACTCAGAAATCGTCTTGTGAATTTGACGCAAGAAAGTGTAAAGCAAGAGAGAGCAGGTC 2077
QY 2069 CTTGTGCTTAACCGGTATCTGTAGATCCACATTTCAAGATTTTGCATGACAGAGATC 2128

Db 2078 CCTTGTGCTTAAGCTGATCTGATAGATCCACATTTTCACGATTTGGTTCGAGATC 2137
OY 2129 TCAAGACACGACGACGACGACCTCACAAAAGTCCCAACCATCGGAGCTATGGAGGCCAG 2188
Db 2138 TCAAGACACGACGACGACGACCTCACAAAAGTCCCAACCATCGGAGCTATGGAGGCCAG 2197
OY 2189 GTTCAGGTAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAGTCTGTTGTGAGTG 2248
Db 2198 GTTCAGGTAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAGTCTGTTGTGAGTG 2257
OY 2249 CGAAGAGGAAACTGCGGCAAAATCATAGGAGTGTAAAGAGATGAGATGATGATG 2308
Db 2258 CGAAGAGGAAACTGCGGCAAAATCATAGGAGTGTAAAGAGATGAGATGATGATG 2317
OY 2309 TTGCTGCTAGAGCTGTGATGATGCTTCTAAATGGGGTTAAGACCCCGTTAACAATC 2368
Db 2318 TTGCTGCTAGAGCTGTGATGATGCTTCTAAATGGGGTTAAGACCCCGTTAACAATC 2377
OY 2369 TGTACATTGATGAGCAATTTGCTGCTCATGACGAGGACGCTGCTGCACTGATTTGCAATC 2428
Db 2378 TGTACATTGATGAGCAATTTGCTGCTCATGACGAGGACGCTGCTGCACTGATTTGCAATC 2437
OY 2429 TCAAACTTAAGAAAGTGTATTGTGCGGGGACCCAAACAATGCGGCTTTTAAACATGA 2488
Db 2438 TCAAACTTAAGAAAGTGTATTGTGCGGGGACCCAAACAATGCGGCTTTTAAACATGA 2497
OY 2489 TGTGCTGAAATGATATTTAATCAATGATGATGATGATGATGATGATGATGATGATG 2548
Db 2498 TGTGCTGAAATGATATTTAATCAATGATGATGATGATGATGATGATGATGATGATG 2557
OY 2549 CTAGAGGATGACACAGACTGTATACCGCATGCTCCACGCTCTTCTTACGACAAAGC 2608
Db 2558 CTAGAGGATGACACAGACTGTATACCGCATGCTCCACGCTCTTCTTACGACAAAGC 2617
OY 2609 TGAAGACGCTTAACCATGTGCTGATTAATCATATAGATACCAAGGACCAACAAC 2668
Db 2618 TGAAGACGCTTAACCATGTGCTGATTAATCATATAGATACCAAGGACCAACAAC 2677
OY 2669 CGCAAAAGATATCTGATTTCAACCTGTTCAAGAGATGGGTGAACAGCTACAGATTG 2728
Db 2678 CGCAAAAGATATCTGATTTCAACCTGTTCAAGAGATGGGTGAACAGCTACAGATTG 2737
OY 2729 ACTACAAAATACGAAATCATGCTGCGGCTGCTGCAAGACTTACGCGAAAGGCG 2788
Db 2738 ACTACAAAATACGAAATCATGCTGCGGCTGCTGCAAGACTTACGCGAAAGGCG 2797
OY 2789 TTTATGCTGTCAAGTAAAGTCAAGAAATCACTCTACTCGAGACTTCTGAGACG 2848
Db 2798 TTTATGCTGTCAAGTAAAGTCAAGAAATCACTCTACTCGAGACTTCTGAGACG 2857
OY 2849 TGAAGCTGTACTTACGACACAGAAAGCGCATTTGTCTGAAGACGCTGCTGATC 2908
Db 2858 TGAAGCTGTACTTACGACACAGAAAGCGCATTTGTCTGAAGACGCTGCTGATC 2917
OY 2909 CCTGATTAAGACATTAAGCTTAATATCCCGGGGATTTCAAGCTTCAATGAGACG 2968
Db 2918 CCTGATTAAGACATTAAGCTTAATATCCCGGGGATTTCAAGCTTCAATGAGACG 2977
OY 2969 GGCACGCGAAGACGACGCTTATGACGCGTCTTGTATAGCGACGACGCTGATG 3028
Db 2978 GGCACGCGAAGACGACGCTTATGACGCGTCTTGTATAGCGACGACGCTGATG 3037
OY 3029 TGTTCAGATTAAGTGAAGCTGTGCTGGCGAAGGCTTTAAGACGCTTGGCCACG 3088
Db 3038 TGTTCAGATTAAGTGAAGCTGTGCTGGCGAAGGCTTTAAGACGCTTGGCCACG 3097
OY 3089 CCAACATTTGCTGACGAGACGAGTGGGAGACGTTGACCCATTTCAAGACGACG 3148
Db 3098 CCAACATTTGCTGACGAGACGAGTGGGAGACGTTGACCCATTTCAAGACGACG 3157
OY 3149 CCACTCACTGAAATGGCACTGAACTTTCTTTGACCAAGCTTCTTTGAGTAGACCTG 3208

Db 3158 CCACTCACTGAAATGGCACTGAACTTTCTTTGACCAAGCTTCTTTGGTGAACCTG 3217
OY 3209 ACAAGGGTATATTTCCGCTCTTACCGATGACATTAAGAGGATGACACTGGAT 3268
Db 3218 ACAAGGGTATATTTCCGCTCTTACCGATGACATTAAGAGGATGACACTGGAT 3277
OY 3269 ACTGCGAGGAGAAATGATATGGGCTTAATAGAGGTAGCAAGAGATTTGTCAGCG 3328
Db 3278 ACTGCGAGGAGAAATGATATGGGCTTAATAGAGGTAGCAAGAGATTTGTCAGCG 3337
OY 3329 GATATCCGTGATCACAAAAGCGTTGACACAGGAGGATGATATTAAGATTAAT 3388
Db 3338 GATATCCGTGATCACAAAAGCGTTGACACAGGAGGATGATATTAAGATTAAT 3397
OY 3389 CCATCAAGACATCTCTCAAAATTAATGATGCTTCAATTAATGCGGCTTCCCACT 3448
Db 3398 CCATCAAGACATCTCTCAAAATTAATGATGCTTCAATTAATGCGGCTTCCCACT 3457
OY 3449 CGTTGATGTTGACCAAGAGACAGGATACATGATCAAGCGGATTTCTATTAAGA 3508
Db 3458 CGTTGATGTTGACCAAGAGACAGGATACATGATCAAGCGGATTTCTATTAAGA 3517
OY 3509 TGAAGGCAATCTGTGTTGGTATCGCGATCTTATGACATTCAGGAAAGATG 3568
Db 3518 TGAATGCAAAATCTGTGTTGGTATCGCGATCTTATGACATTCAGGAAAGATG 3577
OY 3569 AGTCATGGGTCCATGCGCACTAATACATCAGGTGATGATGATGATGATGATGATG 3628
Db 3578 AGTCATGGGTCCATGCGCACTAATACATCAGGTGATGATGATGATGATGATGATG 3637
OY 3629 GGCATGCGTAAATATGACATTAATCTTTGCAATGTTAGACCCCGTACAGAAACAT 3688
Db 3638 GGCATGCGTAAATATGACATTAATCTTTGCAATGTTAGACCCCGTACAGAAACAT 3697
OY 3689 ACTACAAAGTGGAGATACGCTATACACACAGATGCTTAAAGCTGCTG 3748
Db 3698 ACTACAAAGTGGAGATACGCTATACACACAGATGCTTAAAGCTGCTG 3757
OY 3749 ACCACTGAACATGCGGGAACATGCTGCTGATGATGATGATGATGATGATGATG 3808
Db 3758 ACCACTGAACATGCGGGAACATGCTGCTGATGATGATGATGATGATGATGATG 3817
OY 3809 CCGAATATATCATCTCGGTGACCGCTATTTAGGTTTAAACCGTCTGTGACCTA 3868
Db 3818 CCGAATATATCATCTCGGTGACCGCTATTTAGGTTTAAACCGTCTGTGACCTA 3877
OY 3869 AGAACCTGCGAAATATGAGGTTCTTTGCTTCTGCGCAAGACCAAGCAAC 3928
Db 3878 AGAACCTGCGAAATATGAGGTTCTTTGCTTCTGCGCAAGACCAAGCAAC 3937
OY 3929 ACAACATGACACGACAGATCGGTGATGCTTGAACAATCTATCAAGGTCACCA 3988
Db 3938 ACAACATGACACGACAGATCGGTGATGCTTGAACAATCTATCAAGGTCACCA 3997
OY 3989 GGTACGAGGAGGAGACCTCAGGCTACAGATGATCAGAGGTACATTAAGCAAGG 4048
Db 3998 GGTACGAGGAGGAGACCTCAGGCTACAGATGATCAGAGGTACATTAAGCAAGG 4057
OY 4049 CTGACCAAGCTATCTTAATGCTGCTAATAGCAAGTCAACAGGTTCCGAGTGTG 4108
Db 4058 CTGACCAAGCTATCTTAATGCTGCTAATAGCAAGTCAACAGGTTCCGAGTGTG 4117
OY 4109 GTGACCTGATCCGAAATGCGCGCTGCTTTGATAGACAGCAATAGCTGCGGACG 4168
Db 4118 GTGACCTGATCCGAAATGCGCGCTGCTTTGATAGACAGCAATAGCTGCGGACG 4177
OY 4169 CTAGACTGTGAGACGAAACGCTCATATACATGCTGTAGAGACCAATTTTGTAGA 4228
Db 4178 CTAGACTGTGAGACGAAACGCTCATATACATGCTGTAGAGACCAATTTTGTAGA 4237
OY 4229 TGCAGGAACGAGAGGACCTTAAGCTGCAAGCTGCTACATGAGCATAGCTGATC 4288
Db 4238 TGCAGGAACGAGAGGACCTTAAGCTGCAAGCTGCTACATGAGCATAGCTGATC 4297

OY	4289	CCAAACGCTGACGGATTCCAAATAATAGTACCGCTACCTGCAACCGGCATCTATTCTG	43348
Db	4288	TCMAACGCTGACGGATTACTTAATATACGACCGCTACTCTGCACTGCACTTATTCTG	4357
OY	4349	GTGGCAAGATCGAGTGAATGCATCTTGCATCACTCTGTTCAGCTCTTCGACACTACG	4408
Db	4358	GTGGCAAGATCGAGTGAATGCATCTTGCATCACTCTGTTCAGCTCTTCGACACTACG	4417
OY	4489	ATGCCAGATGACCACTAATTGCTTGGATTAACAATGGGACACAGATTAATCGAGGCCA	4468
Db	4418	ATGCCAGATGACCACTAATTGCTTGGATTAACAATGGGACACAGATTAATCGAGGCCA	4477
OY	4469	TTCAACCGCAAGAAAGCGTCGAAATTCG---GATGATGACAAGCCAGTACATTCGACT	4525
Db	4478	TTCAACCGCAAGAAAGCGTCGAAATTCGATGATGATGATGACAGGCACTACATTCGACT	4537
OY	4526	TGCTCAGGGTCCACCCAAACAGCTCTTTGGCAGGACGACAGGTTTACTCGCTCAATAGG	4585
Db	4538	TGCTCAGGGTCCACCCAAACAGCTCTTTGGCAGGACGACAGGTTTACTCGCTCAATAGG	4597
OY	4586	GCAAGTTGTATTCACTCGAAGGTACACGATTCATCAACCGCCAAAGACATTTGCG	4645
Db	4588	GCAAGTTGTATTCACTCGAAGGTACACGATTCATCAACCGCCAAAGACATTTGCG	4657
OY	4646	AAATTCATGCAATGTGGCCCAACAAATCTGAGGCTTAATGACAGATTTGCTTTGATCATCC	4705
Db	4658	AAATTCATGCAATGTGGCCCAACAAATCTGAGGCTTAATGACAGATTTGCTTTGATCATCC	4717
OY	4706	TGGGGGAAGATATGTCGACGACATCGCTCCAAATGCCGATGAGGAGTCAGAGCGCTCG	4765
Db	4718	TGGGGGAAGATATGTCGACGACATCGCTCCAAATGCCGATGAGGAGTCAGAGCGCGTCTG	4777
OY	4766	CTCCACCTCACAACCTTCATGCTCGTGTGTAATTACGTATGACAGGCTGAGCGGTATCA	4825
Db	4778	CTCCACCTCACAACCTTCATGCTCGTGTGTAATTACGTATGACAGGCTGAGCGGTATCA	4837
OY	4826	GCTTGCCGCTTCGCCAAGAAAGAACAGTTTCGCCGTATGCTCATCACTTCTGTTGCCGAAGT	4885
Db	4838	GCTTGCCGCTTCGCCAAGAAAGAACAGTTTCGCCGTATGCTCATCACTTCCGTTGCCGAAGT	4897
OY	4886	ACAGGATCACAGGCGTGCAGAAAGTACAGTGCAGCAACACAGTCTGTTTTCAGGCGTGC	4945
Db	4898	ACAGGATCACAGGCGTGCAGAAAGTACAGTGCAGCAACACAGTCTGTTTTCAGGCGTGC	4957
OY	4946	TACCAACGCGCTGTACACCCACAGAAAGTACGCGAAATATTCTTGAAACGCCACC	5000
Db	4958	TACCAACGCGCTGTACACCCACAGAAAGTACGCGAAATATTCTTGAAACGCCACC	5012

RESULT 2

```

US-09-454-721A-3
: Sequence 3, Application US/09454721A
: Patent No. 6296854
: GENERAL INFORMATION:
: APPLICANT: Peter Pushko
: APPLICANT: Michael D. Parker
: APPLICANT: Jonathan F. Smith
: APPLICANT: Bruce J. Crise
: TITLE OF INVENTION: Live Attenuated Venezuelan
: FILE REFERENCE: Amy 146
: CURRENT APPLICATION NUMBER: US/09/454,721A
: CURRENT FILING DATE: 1999-12-07
: PRIOR APPLICATION NUMBER: US 60/111,330
: PRIOR FILING DATE: 1998-12-07
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: Microsoft Word
: SEQ ID NO 3
: LENGTH: 11459
: TYPE: DNA
: ORGANISM: Venezuelan Equine Encephalitis Virus
: FEATURE:
US-09-454-721A-3

```

Query Match	42.4%;	Score 218;	DB 3;	Length 11459;
Best Local Similarity	64.5%;	Pred. No. 0;		
Matches 3219;	Conservative	0;	Mismatches 1745;	Indels 24;
				Gaps 3;

QY	10	ACTATGATCCAAATATGGAAGAATTACGTTGACTTAAGATGTCGACGCGCCGATATGC	69
Db	30	AATTAACCTACCCMAAATGAGAAAGTTCACTTTGACTGAGGAAGAACGCCCATTTCCCTC	89
QY	70	AAGTCGTTACGCGGACGTTTCCAAATTTGATGCGAAGGACGAGTCACTGACAAAT	129
Db	90	AGAGCTTTGACCGGAGCTTCCCGCAGTTTGAGTAGAAGCCAAAGAGTCACTGATTAAT	149
QY	130	GACCATGCGCAATGCCAAGCGCTTTTCGATGTGGCACAAGGCTCATTTGAAGCGAATTC	189
Db	150	GACCATGCTAATGCGACAGCGCTTTTCGATGTGCTTCMAAATCATTCGAAACGGAAGGTG	209
QY	190	GACCGGGACCAAGTTATCTTGACATTTGGAAGTGGCGCCGTCAGACATGCAATTCCAAT	249
Db	210	GACCCATCCGACAGATCTTGTAATTTGGAAGTGGCGCCCGCAGATATGATTTCTAAG	269
QY	250	CACCGCTATCATTTGATCTGCGCTATGATTAAGCGCTGAAGA CCGCGACAGACTACACAGG	309
Db	270	CACAAGTATCATTTGATCTGTCGATGAGATGTGGGAAGATCCGGA CAGATTGTAATAAG	329
QY	310	TATGCAAGAAACCTTAAGAAAAGT-----GACATTACCGACAGAAACATAGCCTCTAAG	363
Db	330	TATGCAACTAAGCTGAAGAAAAACGTGAAGAAATTAATGATTAAGAAATTTGACAAGAA	389
QY	364	GCGGACACTGTGGAAGTCATGTCAACACCAAGCAGAGACTCCATCTGTGTAATG	423
Db	390	ATGAAGAGAGCTTCGCGCGCCGCTATGAGGACCCCTGACCTGGAACATGACATATGTGCTC	449
QY	424	CACA CAGACGCGACGTTGAGTACTTTGGAAGTGAAGTACGATTAACCAAGATGTGTAGCA	483
Db	450	CACGACACGAGTGTGTGCGCTTACGAAAGGCGAAGTCGTGTTTACAGAGATGTATACGCG	509
QY	484	GTCCATGACCGGACATCAATCTTACCA CCAAGCGCTTAAGGAAGTTAGACAAATTTACTGG	543
Db	510	GTTGACGACCGGACAAAGTCTCTATACCCAAACCAATTAAGGAAGTTAGTGCCTCACTGG	569
QY	544	ATTAGGCTTTGACAGACCCCTTTTATGTACA AAAACATGAGAGGTTCCTACCTCACTTAC	603
Db	570	ATTAGCTTTGACACACCCCTTTTATGTGTTTAAGAACTTTGGTGAGCATATCAATCATAC	629
QY	604	AACA CGAACCTGGGCTGACGAGAGAGTATGGAACA CATTGCTCGTGAATCTCA	663
Db	630	TCTACCAACTGGGCGACGAAACCGTGTAA CCGCTCGTAA CATAAGGCTCTATGCAAGCTCT	689
QY	664	GATCTTCAGGAGACAGGCTTGGA AAACCTTCATCCCTTAGGAAGAAGGCTCCAACCT	723
Db	690	GACGTTATGAGCGGTACGTAAGAGGATGTCCATCTTAAGAAAGAAATTTGAAACCA	749
QY	724	ACTAATAAGATCATATTTCTCGGTGGTTTCAACAATCTACACAGAAGATAGATCACTGTTA	783
Db	750	TCCAACAATGTTCTATTTCTGTTGGCTCGACATCTACACAGAAAGAGGGAATTAATG	809
QY	784	CGTAGCTGGACCTTCCAAACGTGTTCAC TTTGAAGAAAGTCTAACTTCA CAGATAGA	843
Db	810	AGGAGCTGGACCTCGCGTGTGTAATTTCACTTA CGTGGCAAGCAAAAATTACATATGTCGG	869
QY	844	TGTGGGACCAATGTCA GCTGTGAAGGGAAGCTAATCAAAAAGATATAGATAGCGCCAGGA	903
Db	870	TGTAGATATATATGTTAGTTAGTGCAGCGGTA CGTGTTAAACAATAAGCTATATAGTCCAGGC	929
QY	904	CTATACGTTAAGTTGAGAACTTGGCGTCCACAATGCA TCGAGAGGTTTCTTGAGTTGC	963
Db	930	CTGTATGGAAGGCTTACAGGCTATAGCTGTACATGCA CCGCGAGGGAATTTGTGTGCTGC	989
QY	964	AAAGTCACAGATACGCTGCGCGGCGAAGAGGTTTCTTTGCTGTGTGTATGTAATGTACCA	1023
Db	990	AAAGTGCACAGCAATTTGAACGGGGAGAGGGTCTCTTTTCCCGTGTGACATATAGTGC	1049

Qy 1024 GCCACACTTTGCGATGAGATGACAGGGAATTCGCGACATGACGTATGATGATGACGCA 1083
Db 1050 GCTACATGTTGTGACCAATATGACTGGCATACAGAGTGTCAAGTGGCGGACGACGCG 1109
Qy 1084 CAAAACCTATTTGTTGGGCTCAACCAAGATTTGCTCATGTTGATGAGACGCAAGAAAT 1143
Db 1110 CAAAAACGCTGTTGGGCTCAACGAGGTATGATGCTCAACGATGCGACCCAGGAAAC 1169
Qy 1144 ACTAACACATGACAGATCTATTTACAGTGTGCGCGAGGCGTTTCCAGGTTGGGCG 1203
Db 1170 ACCAATACATGAAAAATTTACCTTTTGCCGATGAGGCGCAGGCAATTTGCTAGTGGGCA 1229
Qy 1204 CGTGAACATCGTGCAGCTTGGACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTCTT 1263
Db 1230 AAGGATTTTAAGGAAGATCAAGAGATGAAAGGCCACTAGACTACGAGATAGACATTA 1289
Qy 1264 ACTATGGGCTGCTGTTGGGCTTTCAAGACCCAGAAATCAATCATCTTACAGAAAGCT 1323
Db 1290 GTCAATGGGCTGTTGGGCTTTTGAAGGCACAAGATACATCTATTTATTAAGGCGCG 1349
Qy 1324 GGTAAGCAAAATTAAGAAAGTACCTGCGCTTGTGCTCATTTGTGATTCAGGCTT 1383
Db 1350 GATACCCAAACCATATCAAAAGTGAACAGGATTTCCACTCATTTGCTGCTCCAGATA 1409
Qy 1384 ACCAGCCAAGGAGCTCGATATGAGGCTTCCGCGTAGGCTCAAGCTGCTGTAACCAACT 1443
Db 1410 GGCAGTACACATTTGAGATGGGGCTGAGAACAAAGATCAGGAAATTTTGAAGAGACAC 1469
Qy 1444 GTCAAAACCCGACCGGCTATTTCAATGCGCCGATGTGAGCATCTGCGTGGCTTACAGAA 1503
Db 1470 AAGGAGCCGTCACCTCTCATTAACCCGAGAGCATCAAGAAAGTAAAGTGGCGACCCAT 1529
Qy 1504 GAAGCTGAAGAGTGGCTGACGCGGAGAGATCAAGAAAGCCCTCCACCTTGTCTCT 1563
Db 1530 GAGGCTAAGAGAGTGGCTGAAGCCGAGAGTTGGCGGAGCTTACCACTTTTGGCAGCT 1589
Qy 1564 GAAATAGAAAAAGAGCCGTAGAGGCAAGAGTGAAGCTCATATATCAAGAGGACAGACA 1623
Db 1590 GATGTGAGGAGCCCACTCTGGAACCGATGTGACTTGAATGTTCAAGAGGCTGGGCGC 1649
Qy 1624 GGTAGCTGAGACACACAGAGACACATCAGGGTGAACAAGTTACCAGCGGAGAGAAAG 1683
Db 1650 GGCCTACGTGAGACACCTGCTGGCTTGAATTAAGTTACAGCTACGCTGGCGAGAGACAAG 1709
Qy 1684 ATTGGGCTTACGCTATATCTTTACCCAGCGGATTTGAATGTAAGAAATCTGCGCT 1743
Db 1710 ATCGCTCTTACGCTGTCTTTCTCCGAGGCTTACTCAAGAGTAAAAATTAATCTTGC 1769
Qy 1744 ATCCACCCATTTGCGGAGCAAGATTAATGATTAATGATCAAAAGTAAAGGAGGAGATAC 1803
Db 1770 ATCCACCTCTCTGCTGAACAACTCATAGTATACACACTCTGCGCCGAAAAAGGCGTTAT 1829
Qy 1804 AAAGTCAGCATACACGCTAAGGTCAATTTGACAGAAAGGACGCGGCTCCCTGTTCA 1863
Db 1830 GCCCGTGAACCATACATGATTAAGTATGTTGTCAGAGGACATGCAATACCCGTCAG 1889
Qy 1864 GACTTCCAGGATTAAGTGAAGGCGCTACGATCGTTTCAACGAGAGGAGTTGTTAAAC 1923
Db 1890 GACTTCAAGCTCTGAAGTGAAGTGCACATTTGTATCAACGAACTGAGTTCCTTAAC 1949
Qy 1924 AGATACCTGACACATGCAATCAACGAGAGGCGCTAAACACTGAGGAAAGTACTAT 1983
Db 1950 AGGTAAGCTGACACATTTGCTGACACATGAGAGGCGCTGAACTGTAAGAAATATTAAC 2009
Qy 1984 AAGACTGTAAGACTCAGAGACACAGACTCAGAAATCGTCTTGTATATTGAGCGACGAAG 2043
Db 2010 AAAACTGTCAAGCCAGAGAGACAGAGCGGAATCTGTATGACATGACAGAAACAG 2069
Qy 2044 TGTGTAAAGCGAAGAGAGGAGTCCCTTTGTGCTTAAACCGGTGATCTGTATGATCAACA 2103
Db 2070 TGCCTCAAGAAAGAACTAGTCACTGGGCTAAGGCTCAACAGGAGACTGTGATCTCTCC 2129
Qy 2104 TTTCAAGATTGCGTACGAGAGTCTCAAGACACAGCAGCAGCTCACAAAGTCCCA 2163

Db 2130 TTCCATGAATTCGCTTACGAGAGTCTGAGAACAGACAGCGCGCTTACCAATACCA 2189
Qy 2164 ACCATCGAGTCTATGAGAGTCCAGGTTCAAGTAATTTGAAATCATCAAAAGCCTGTG 2223
Db 2190 ACCATAGGGGTATGAGGTTCCAGAGATCAGCAAGTCTGCAATCAATTAAGCCCACTC 2249
Qy 2224 ACTAAGAAAGATCTGTTGTGATGTCGAAAGGAAACCTGCGCAAAATCATCAGGAT 2283
Db 2250 ACCAAAAAGATCTAGTGTGATGAGGCGCAAGAAAGAAATCTGCAAAATTTAAGGAC 2309
Qy 2284 GTAAGAGATGAGACGTATGATGTTGCTCTGAGACTGTCGATTCAGTCTTAAAT 2343
Db 2310 GTCAAGAAAAATGAAGGCTGAGCGTCAATCCAGAACTGTGAGCTGAGTCTTGAAT 2369
Qy 2344 GGGGTTAAGCACCCCGTTTAACTCTGTATCAATTGATGAGGCAATTTGCTGCCATGAGG 2403
Db 2370 GGAATGCAAAACCCCGTATGAGACCCCTGTATTTGACGAAGCTTTGCTGTATCAGAGT 2429
Qy 2404 ACGCTGCGGACATGATTGCGATGCTCAACCTAAGAAAGTATGTTGGGGGAGCCCA 2463
Db 2430 ACTTCAGAGCGCTCATACCATTTAATGACCTTAAAGGCAAGTCTTGGGAGATCC 2489
Qy 2464 AAACATGCGGCTTCTTTAATGATGTCCTGAAAGTACATTTTAACATGACATATG 2523
Db 2490 AAACATGCGGCTTCTTTAATGATGTCCTGAAAGTACATTTTAACATGAGATTTGC 2549
Qy 2524 ACTGAAGTATCAATTAAGCATCTTGAAGTGTGACACAGACTGTAAACCGCATGCTC 2583
Db 2550 ACACAAAGCTTCCCAAAAGCATCTCTGCGGTGACATTAATGTGATCTTGGCTGCTC 2609
Qy 2584 TCCAGCGCTCTTACGACAAAGCATGAGAGCGGTTAACCATGTCGATTAATATATC 2643
Db 2610 TCAACCTTGTTTACGAAAAAATGAAACGAGATCCGAAAGACATTAAGATTTGTG 2669
Qy 2644 ATAGATACACAGGAGCCCAAAAGCCGACAAAGTATGATTTCTAATCTGTTTCA 2703
Db 2670 ATTGACACTACCGGAGTACCAAACTTAAGAGAGGACATCTCATCTTCACTTTTCA 2729
Qy 2704 GGAATGGGTGAACAGCTACAGATTTGATCAAAAAATCAAGAAATCATATGCTGGCTCA 2763
Db 2730 GGGTGGGTGAAGCAGTTGCAAAATGATTAACAAAGCAACGAAATTAATGACCGCAGCTCC 2789
Qy 2764 TCGCAAGGACTTAGCGGAAAGGCTTATGCTGACGTTCAAAAGTAAAGTGAAGAGATCA 2823
Db 2790 TCTAAGGCTGACCCGTAAGGTGTATATCGTTGCTGACAAAGTAAAGTAATCTT 2849
Qy 2824 CTCTACTGACAGACTTCTGACACAGTGAACGTTGTTACTTACAGCAGCAAGAAAAAGCAT 2883
Db 2850 CTGTACGACCCACCTGAGAACATGTGACGTCCTACTGACCCGACGAGAGACCGCATC 2909
Qy 2884 GTCTGGAAGAGCTAGCTGTGATCTCTGATTAAGACATTAACGTTAAATATTCGCGG 2943
Db 2910 GTGTGAAAAACATAGCCGCGGAGCCATGTAAACACATGACCTGCAAGTACCTGGG 2969
Qy 2944 GATTTCACGCTTCAATTGAGACATGCGAGGCGCAACGACGCAATTAATGCGACGCTT 3003
Db 2970 AATTTCATGCGCATGATGAGAGAGTGGCAAGACGATGATGCTATATGAGGACATC 3029
Qy 3004 CTGTATTAAGCCGACAGACAGTGTATGTTCAAGATAAGTGAAGTCTGCTGGCGAAG 3063
Db 3030 TTGAGAGAACCGGACCCCTACGAGAGTCTTCCAGATTAAGCAAGCTGTGTTGGCCAG 3089
Qy 3064 GCTTTAAGCCAGTCTTGCCACGACGCAATTTGCTGACGAGACAGCAGTGGAGAGC 3123
Db 3090 GCTTTAAGTCCGAGTGTGAAGACCGCTGGCATTAACATGACCACTGAACATATGAACT 3149
Qy 3124 TTGACCATTCAGACATGAGAGCGCTACCTGAAATGAGCACTGAATCTTTTGC 3183
Db 3150 GTGATTAATTTTGAACGGAACAAAGCTCACTCAGAGAGATGATTAAGCAACATATAGC 3209
Qy 3184 ACCAGTTCCTTTGAGATGACCTGGAACAGTGGTTAATTTTCGCTCTCAACGTCGACCT 3243

```

Db      3210 GTGAGTCTTCTTGGACTGATCTGAGCTCCGGTCTATTTTGTGACCCACCTGTTCCGTTA 3269
Qy      3244 ACTTACAGGATATCGACTGGGATTAATCCGACAGGAAGAACATGATGGCTTAATAGA 3303
Db      3270 TCCATTGAGATATCACTGGGATTAATCCCGTGGCTTAACATGTAAGGGCTAATPAA 3329
Qy      3304 GAGGTACCAAGAGGTTGTACGGCGGATATCCGTGATCACAAAAGGGGTGACACAGGC 3363
Db      3330 GAATGGTCCGCTGCTCTCCGAGGTACCCACAACTGCTCGGGAGTTGCCACTGGA 3389
Qy      3364 AGGTACCTGATATAGGAATATATCAATCAAGACTACTCTCCAAATTAATGTGTT 3423
Db      3390 AGAGTATGATGACGAACACTGTGTACACTGGCAATTAATGATCCGCGATTAACCTAGTA 3449
Qy      3424 CCATTAATGGCGGTTGCCCACTGTTGATGTTGACCAAGAGCAGAGGTACAACT 3483
Db      3450 CCGTAAACAGAAACTGCTCAATGCTTAATGCTTCCACCAATTAAGAACCCACAGAGT 3509
Qy      3484 GATCACAGCGGATTCCTATCTAAGATGAAGGCAAAATCTGTGTGATGATCGCGATCCT 3543
Db      3510 GACTTTTCTTCACTGCTGACGAATTAAGAGGCAAGAACTGCTGTGTGTTGGGAAAG 3569
Qy      3544 ATCAGCATTCAGGGAAGAAAGTAGAGTCCATGGGTCCATTGCCACTAATATACATCAGG 3603
Db      3570 TTGTCCTCCAGGCAAAATGTGTGACTGTGTGACAGCCGCGCTAGAGCTACCTTGA 3629
Qy      3604 TGTGATCTCAATTTGGGAATCCATAGCCATGTCGTAAATATGACATTAATCTTGTCAAT 3663
Db      3630 GCTCGCTGATTTAGGATCCAGGTGATGTCCAAATTAATGACATTAATTTGTGTAAT 3689
Qy      3664 GTTAGAGACCCCGTACAGGAACATCACTACCAAGTGCAGAGTCAAGCTATCCACAC 3723
Db      3690 GTGAGAGCCCATTAATATATACATCACTATGACAGTGTGAAGACATTCATTAAGCTT 3749
Qy      3724 AGCATGCTAACGTGTAAAGCTGTCCACCACTGAACACTGGCGGAACATGTGTGCTATA 3783
Db      3750 AGCATGTGTACCAAGAAAGCTGTGTGATGATCCCGCGGAACCTGTGTGACGATA 3809
Qy      3784 GGGATATGGGCTTGTGATCCGCGCAACGAAATATCATCACTGGGGTGGACGCTCAAT 3843
Db      3810 GGTATGATGATGCTGACAGGCGCAAGAACATCATTTGGTGTATAGCCGCGAGTTC 3869
Qy      3844 AGGTTATCCCGTGTCTGTCAAGCTTAAGAACACTGCGGAAATTAAGGTTCTGTCTGTG 3903
Db      3870 AAGTTTCCCGGATATGCAACCGAAATCTCACTGTAAGAGAGGAAGTTCTGTGTGTA 3929
Qy      3904 TTCTTCGGAAGAACAAAGCGCAACCAACATGACGAGACAGACTCGGTGTGTGCTT 3963
Db      3930 TTCAATGGGTACGATCCGAAAGGCCGCTGACGCAATCTTCAAGCTTTCATCAACCTTG 3989
Qy      3964 GACAAATCTATCAAGGCTCAACAGATAGAGGCAAGGACGTCACGCTACAGAGTG 4023
Db      3990 ACCAAATTTATACAGGTTCCAGACTCCACAGAGCGGATGTGACCTTCATATCATGTG 4049
Qy      4024 ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCGTTAATGCTTAATAGCAA 4083
Db      4050 GTGGAAGGGGATATTTGCCAGGCGCAACGAGAGATTAATTAATGCTGCTAAACGAAA 4109
Qy      4084 GGTCAACCAAGTTCGCGAGTGTGCGGTGCTGTATCCGAAATAGCGCGGCTCTTTGAT 4143
Db      4110 GGAACAACTGCGGAGGGGTGTGCGGAGCGCTGTATTAAGAAAGTTCCCGAAAGCTTCAT 4169
Qy      4144 AGAAGCAATAGCTGTGCGGACGCTGATTTGTGAAGCACGAAAC---GCTATCTA 4200
Db      4170 TTACAGCGGATCGAAGTGAAGAAAGCGGACTGTGCAAGGTGACGCTTAAACATATCTT 4229
Qy      4201 CATGCTGAGAGCAACCAATTTTCTAAGATGCGGAAACGAGGGGACCTTAAGCTGCA 4260
Db      4230 CATGCCGAGAGCAAACTTCAACAAAGTTTCGAGGTGAAGGTGCAAAACAGTTGCA 4289
Qy      4261 GCTGCTTACATGAGCATAGCTCATGCTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db      4290 GAGGCTTATGAGTCATGCTCAAGATGTCAACGATTAACAAATGAGTAGCGAT 4349

```

```

Qy      4321 CCGCTACTGTCACCCGCGCATCTATTTGTTGGCAAGATGAGTATGATCATTCAT 4380
Db      4350 CCACTGTGTTCACCCGCGCATCTTTTCGGAAGAAAGATGATCAACCCATCATTTAAC 4409
Qy      4381 CACTGTTCACGCTTTTGCACACTACGATGCGGATGTCACATATATGCTTGGATPAA 4440
Db      4410 CATTTGCTGACAGTTTAAACCACTGATGACAGATGATGATCAATATCTCAGGAGCAAG 4469
Qy      4441 CAATGGAGACCAAGATTAATGAGGCTTACCGCAAAAGAAAGCTGCAAAATTTGGA- 4499
Db      4470 AAATGGAAATGACTCTCAAGAAAGAGTGGCTAGAGAGAAAGCAGTGAAGATATGC 4529
Qy      4500 -----TGATGACAAAGCCAGTAGACATTAAGCTTTGGTCAAGGCTCAACCAAC 4545
Db      4530 ATATCCGACAGCTTTTATGATGACGAACCTGATGACAGAGTGTGAGGGTGTATCCGAG 4589
Qy      4546 AGCTCTTTGGCAGGACAGCAGGTTACTCCGTCAATGAGGCAAGTTGATTCATACCTG 4605
Db      4590 AGTTCTTTGGCTGGAAGAAAGGCTACAGCAACAGCGATGGCAAACTTTCTCATATTTG 4649
Qy      4606 GAAGGTACAGATTCATGACCGCCAGAGACATTTCCGAATTCATGCAATGTGCCC 4665
Db      4650 GAAGGACCAAGTTTCAACAGCGGCCAGAGATATGACAAATTAATGCTATGTGCCC 4709
Qy      4666 AACAAATCTGAGGCTAATGAGCAGATTTGCTTGTACATCTCGGGGAGAGATGTCCAGC 4725
Db      4710 GTTGCAAGCGAGGCCATGAGAGTATGATGATATCTCTCGAGAAAGATGATAGAGT 4769
Qy      4726 ATCCGCTCCAATGCCCCAGTAGAGAGTCAAGAGCGTCTGCTCACTCAACACTTCCA 4785
Db      4770 ATTAGTTCGAATGCCCCGCTGAAGAGTCCGAAGCTCCACACCACTAGACCGCTCCT 4829
Qy      4786 TGCTGTGTATTAAGCTATGAGGCTGAGCGCTGATATACAGTGTGCTGTGCAAGAAA 4845
Db      4830 TGCTTGTGATTCATGTCATGATGATCCAGAAAGATGACGCGCTTAAAGCTTCAAGTCA 4889
Qy      4846 GAACAGTTCGCGGATGCTATCATCTCTGTGGCGGAAGTACAGATCAACAGCGGTGAG 4905
Db      4890 GAACAAATTAATGATGATGATCTTTCATTTGCTCAACCGAAAGTGTGCTGTATTCATCA 4949
Qy      4906 AAGCTACAGTGCAGCAAAACAGTCTGTTTTCAGGCGCTGTACCAACGCGCTGTACACCCC 4965
Db      4950 AAGATCCAAATGCTCCAGCTATATGTTCTTCAACCGAAAGTGTGCTGTATTCATCA 5009
Qy      4966 AGGAAGTA 4973
Db      5010 AGGAAGTA 5017

RESULT 3
US-08-991-840A-2
; Sequence 2, Application US/08991840A
; Patent No. 6261570
; GENERAL INFORMATION:
; APPLICANT: Michael D. Parker
; APPLICANT: Jonathan F. Smith
; APPLICANT: Bruce Crise
; APPLICANT: Mark Steve Oberste
; APPLICANT: Shannon Schumura
; TITLE OF INVENTION: Live Attenuated Virus Vaccines for Eastern Equine Encephalitis
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: USA MRC - MRC-JA
; CITY: FORT DETRICK, FREDERICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5

```

SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991.840A
FILING DATE: December 16, 1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Provisional Application 60/047162,
FILING DATE: May 20, 1997
APPLICATION NUMBER: Provisional Application 60/053,652
FILING DATE: July 24, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Charles H. Harris
REGISTRATION NUMBER: 34,616
REFERENCE/DOCKET NUMBER: 003/058/SAP RIID 96-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 619-2065
FAX: (301) 619-5034
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 11464 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
US-08-991-840A-2

Query Match 41.5%; Score 2074.8; DB 3; Length 11464;
Best Local Similarity 64.1%; Pred No. 0;

Matches 3202; Conservative 0; Mismatches 1772; Indels 24; Gaps 4;

QY 1 ACCCTCAAACTAATCGATCCAAATATGGAAGAATTCAAGTTAGTACTGATGCTGACAC 60
DB 20 AGCCCAAAACCAATACTAACCCAAATATGGAAGAATTCAAGTTAGTACTGATGACATAGT 79
QY 61 CCGTATGTCAGTCTGTACAGGAGACGTTTCCACAATTGGAATGGAAGCAAGCAAGTGC 120
DB 80 CCGTATGTCAGTCTGTACAGGAGACGTTTCCACAATTGGAATGGAAGCAAGCAAGTGC 139
QY 121 ACTGCATATGACCATGCAATGCAAGGAGGTTTGGCATGTCAGTCAAGCAAGTCAAGT 180
DB 140 ACAGATATGACCATGCAATGCAAGGAGGTTTGGCATGTCAGTCAAGCAAGTCAAGT 199
QY 181 AGCGAAGTGAACCGGAGCAAGTATCTTGAACATTTGAAGTGCAGCCGTCAGACATCA 240
DB 200 ACGGAGTGAACCGGAGCAAGTATCTTGAACATTTGAAGTGCAGCCGTCAGACATCA 259
QY 241 CATTCATATCAAGGATATCTTGAAGTGCAGCCGTCAGACATCAAGTGCAGACATCA 300
DB 260 TATTCATATCAAGGATATCTTGAAGTGCAGCCGTCAGACATCAAGTGCAGACATCA 319
QY 301 CTACAAAGGATATGCAAGGATATCTTGAAGTGCAGCCGTCAGACATCAAGTGCAGAC 354
DB 320 CTGTTTAAATATGACCAAGGATATCTTGAAGTGCAGCCGTCAGACATCAAGTGCAGAC 379
QY 355 GCGTCTAAGGCGGAGACCTGCTGGAAGTCAATGCAACCGAGAGCAAGTCAAGTCAAGT 414
DB 380 GACAAAGATGGAAGAGGCTTGGAGATGCAAGCAAGTCAAGTCAAGTCAAGTCAAGT 439
QY 415 CTGTATATGACACAGAGGCGGAGTATCTTGAAGTGCAGCCGTCAGACATCAAGT 474
DB 440 ATTTCCTTCAAGAGATGAAGCTGCTGATTTGAAGGCTCAAGTGCAGTATCAAGT 499
QY 475 GTGTAGGAGTCCATGACCGGATCAATCTTGAAGTGCAGCCGTCAGACATCAAGTGCAGAC 534
DB 500 GTGTAGGAGTCCATGACCGGATCAATCTTGAAGTGCAGCCGTCAGACATCAAGTGCAGAC 559
QY 535 ATTACTGATATGCTTGAACAGACCGCTTTATGATCAAAAATGAGAGTCTTCTAC 594
DB 560 GCGTATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 619
QY 595 CCTATCTTCAACAGAACTGGGCTGACAGAGATGATGGAAGCAAGTCAAGTCAAGTCAAGT 654
DB 620 CCGTCTATGTCAGCAAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679

QY 655 GTTACTCATGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
DB 680 TCGAGCTCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739
QY 715 CTCAAGCTTAAATATGATATATCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 774
DB 740 TTAATAACCATCAATTAAGTCTTGTCTTGTAGATCTTACATCTTACAGAGAGAGAG 799
QY 775 TCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 834
DB 800 GACTTACTTAAGAGAGTGGACCTTACCTTCCGTTTTCACCTTACCTTACCTTACCTTAC 859
QY 835 ACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
DB 860 ACTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 919
QY 895 AGCCAGGACTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
DB 920 AGTCCAGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979
QY 955 TTGAGTTCAGAAAGTCAAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1014
DB 980 TTGTGCTGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1039
QY 1015 TATGATCCAGGACCACTTGGCATCAGATGACAGGATTTGCGCAATGAGTATGATG 1074
DB 1040 TACGTGCGCAAGCACTTGTGAGACCAATGACAGGATTTGCGCAATGAGTATGATG 1099
QY 1075 GATGACGCAAAATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134
DB 1100 GACGACGCGCAAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1159
QY 1135 CAAGAATATCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
DB 1160 CAAGAATATCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1219
QY 1195 AGTGTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1254
DB 1220 AGATGCGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
QY 1255 GCGACTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
DB 1280 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1339
QY 1315 AAGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1374
DB 1340 AAGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1399
QY 1375 CCAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
DB 1400 CCAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
QY 1435 GAAACCACTGTCAAACCGGATATTAATGATGATGATGATGATGATGATGATGATGATGAT 1494
DB 1460 GAGGATCTGTCAGACAGACCACTGATTTACCGGCGGATGATGATGATGATGATGATGAT 1519
QY 1495 TTAAGCAAGAGTGAAGAGTGTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1554
DB 1520 GCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
QY 1555 TTGCTCCCTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1614
DB 1580 TTGCTCCCTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1639
QY 1615 GCAAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1674
DB 1640 GCGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1699
QY 1675 GAAAGAAATTTGGCTTACGATATCTTCAACCGGAGGATGATGATGATGATGATGATGAT 1734
DB 1700 GAAAGAAATTTGGCTTACGATATCTTCAACCGGAGGATGATGATGATGATGATGATGAT 1759
QY 1735 CTGCGGTATATCAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1794

Db 1760 CTGACGTGATCCACCCGCTTGACAGAGCAAGTAATGTAATCAACACTCTGGAGAGAG 1819
Qy 1795 GGGAGATATCAAAAGTCGAGCCATACCAAGGTAAAGTCAATTGTAACGAAGAGGACGGCGTC 1854
Db 1820 GGTGATATACGAGTGTGAGCCCTTACCAAGAAAGTGTAGTACCCGAAGGGCAAGCTATA 1879
Qy 1855 CCTGTTCAAGACTTCCAGGCAATGTAGTGAAGCGCTACGATCGTTTTCAAGAGGGAG 1914
Db 1880 CTTGTCCAAAGCTTCCAGAGCACTCAGTAGAGCCGCCAAATTGTGTACAGAAAGTGTAG 1939
Qy 1915 TTGCTAAACAGATACCTGACACCAATGCAATCAACGAGGAGCGCTAAACACTGACGAA 1974
Db 1940 TTCTGTAATAGTATTATACCAACATTTGCAACGACGCTGAGCTGTGAACACAGACGAG 1999
Qy 1975 GAGTACTATAGACTGTAAAGACTCAAGACACAGACTCAAGTAATGCTTGTGATTTGAC 2034
Db 2000 GAGTATTAACAGAGTGTGAAACCTAGTGAACATGAAAGGTGAGTACTTATATGACATTGAC 2059
Qy 2035 GCACGAAAGTGTGTTAAGCGAAGACGACAGTCCCTTGTGCTTAACCGGTGATCTGGTA 2094
Db 2060 AAGAAACAATGTGTGAAGAAAGAACTGTTACGGGACTAGTTTGAACAGAGAACTTGTTC 2119
Qy 2095 GATCCACCATTTCAAGATTGCGTACAGAGTCTCAAGACAGACGACGACGACTTAC 2154
Db 2120 GATCCCCCTTCATGATTTGCAATGAAAGCTTGAAGGACACGTCGGCCGCACTTAT 2179
Qy 2155 AAAATCCCAACATCCGAGTCTATGAGTGTGCAAGTTCAGTAAATCTGGAATCATCAA 2214
Db 2180 CAATATCCAACTATAGTGTGTATGCGTTCGCGGTCCGGAAAGTCTGGAATTAATGAG 2239
Qy 2215 AGCGCTGTGACTAAGAAAGATCTGTGTGAGTSCGAAGAAAGAAACTGCGAGAAATC 2274
Db 2240 AGCGCACTCAAAAGAGAAAGTGTGTGAGTGTGCTAAGAAAGAAAGTGTGCGGAGATA 2299
Qy 2275 ATCAGGATGTAGAGAGATGAGAGTATGAGATTTGCTGTAGAGCTGTGATTCAGTG 2334
Db 2300 ATAAAGGACGTCAAGAGATGAAGGCTGACGCTCAATGCCCCGAGCGGTGACTCAGTG 2359
Qy 2335 CTTTAAATGGGGTTAAGCAACCCCGTTAAACCTCTGTAATTTGATGAGCAATTTGCTGTC 2394
Db 2360 CTGTTAAATGAGATCAAGCAACCCGCTGAGACTTTGTATCATGTATGAAAGCCCTTTCATGTC 2419
Qy 2395 CATSCAGGAGCGTGTGCGCACTGATTGCGCATGCTCAAACTTAAGAAAGTGTATGTC 2454
Db 2420 CAGCCGCGCACTCTCAGGCGCTTAATAGCACTTATAGCCCAAGAAAGCAAGTGTATGT 2479
Qy 2455 GGGGACCCAAACAAATGCGGCTTCTTAAACATGATGTGCTGAAGTACATTTTAAACAT 2514
Db 2480 GGTGACCCAAAGCAATGTGCTTCTTCAATGATGTGCTGAAGTGCATTTCAACCAT 2539
Qy 2515 GACATATGCACTGAAGTATACATTAAGAAAGCATCTCTAGAGAGTGCACACAGACTGTAAC 2574
Db 2540 GAATATATGCACTCAAGTTTCTCAAAAGCAATATCTGCAAGTGTACCAATCTGATGAG 2599
Qy 2575 GCACTGCTTCCAGCTCTTCTAGCAAGAGCAATGAAGCGTTAAACCCATGCTGAT 2634
Db 2600 TCGGTATGTCTCAACCTTTTATGACAAAGAAATGAGAACTACTAACCCCAAGGATTC 2659
Qy 2635 AAAATCATATGATATACCAAGGACCAAAAGCGCAAAAGATGATCTGATTTCAACC 2694
Db 2660 AAAATCAAAATTGACAAACAGAAAGTACTAAACCAAAAAAGATGACTGATTTCTCA 2719
Qy 2695 TGTTCAGAGATGTGGTGAACAGCTACAGATTTGATCTACAAATATCAAGAAATATGACT 2754
Db 2720 TGTTTTGGGGATGGTTTGAACAATGCAAAATAGACTCAAAAGAAATGAAATATATGCC 2779
Qy 2755 GCGGCTCATCGCAAGACTTACGCGGAAAGCGCTTATGCTGAGTACAAAGTCAAC 2814
Db 2780 GCGGCTGCTCAAGGAGTTTACCGGAAAGGTCTATGCACTTAAAGTATCAAAAGTTTAAAC 2839
Qy 2815 GAGATCACTTAACTGCAAGACTTCTGAGCAGCTGAAGCTGTATCTTAAACGACAGAA 2874

Db 2840 GAGAACCAATTTATGCAACCCACACTCAGACATGTAAATGTGCTTGTGACCCGACGAA 2899
Qy 2875 AAAAGCATGTGTGGAAGACGCTTACGTGTGATCCCGGATTAAGACATTTACAGCTAAA 2934
Db 2900 GACAAAGTGTGTGAAGACTCTTTCAGAGGACCCGTGTGATTAAGACCTGACCGCAAG 2959
Qy 2935 TATCCCGGAGATTTCAAGGCTTCAATGAGACTGTGCGAGCGCAACAGACGCTATTAT 2994
Db 2960 TACCCGGAATTTCACTGCAACAATGGAAGATGGAAGGCAAGAAATGATGCATCATG 3019
Qy 2995 GCACGCTTCTTGTATTAACCGCAGACAGCTGATGTGTTCCAGAAATTAAGTGAAGTCTGC 3054
Db 3020 AGACACATCTGTGAGAAACCGGATCCCAAGGATGTCTTCCAAATTAAGCTTAATGTTGC 3079
Qy 3055 TGGCGAAGGCTTTAGGACAGTCTTGGCCACGGCCCAATTTGCTGACGAGACAGACG 3114
Db 3080 TGGCAAAAGGCACTTGATCTGTGCTTAAAGACGCGGGATGATTTGACCAAGACGACG 3139
Qy 3115 TGGGAGAGCTTGACCCCATTTCAAGCATGACAGAGCTACTCACCTGAATGSCACTGAAC 3174
Db 3140 TGGAAACAGTGTGATTTCTCAAGAGGATTAAGGCCCACTCAGTGAAGTGTCTGTAAT 3199
Qy 3175 TTTCTTGCACAGGTTCTTTGAGTAGACCTGACAGTGGGTTATTTTCGCTTCAAC 3234
Db 3200 CAGCTGTGCGTGGCATTTCTCGGTCTAGACTTATGATTTCTGTTTCGCCCCACA 3259
Qy 3235 GTGCACTTCTTAACAGGATTCAGCACTGGGATTAACCTGCGAGGAAAGAAATGTATGG 3294
Db 3260 GTTCACCTCTCAATTAGAAACAACATTTGGGACACCTACCGTCAACCAACATGTAGGG 3319
Qy 3295 CTTAATAGAGGTAGCAAGAGAGATTGTCAAGCGGATATCCGTGATCAGAAAAAGCGTT 3354
Db 3320 TTGAATCAABAAGTGTGACAGCAACTATCAAGAGATACCTCAATTAACACGCGGGTG 3379
Qy 3355 GACACAGGAGGTGACTGATATTAAGGAATTAATCAATCAAGACTTCTCCAACAATT 3414
Db 3380 ACTACTGGAGAGATATGCAATGAACACCGGATCTTTGCGAATTTGATTTCCGCGATA 3439
Qy 3415 AATGTGTTCCATTAATATGCGCGTTGCCCACTCGTTGATTCGTGACCAAAAGGACAG 3474
Db 3440 AATTTATACCGGGGAACCGTCTGCTACACATGCTCTGTGACGAACATGCTGATCAT 3499
Qy 3475 GGTACAACTGATCAACGCGGATTTCTTATAGATGAAGGCAAAATCTGTGTGATGATC 3534
Db 3500 CTTCCAGTATTTTTCGCGCTTTGTGCAAGTGTAAAGGCAAGACGCTCTAGTATGT 3559
Qy 3535 GCGGATCTTATCAGACTTCAAGGAGAAAGTATGATCATGTGCTCATTTGCCACTAAT 3594
Db 3560 GGTGAGAAAGTATGATTTTCAAGGTAAAGACGCTGATGCTTATCTGAAACACCTGATTC 3619
Qy 3595 ACCATCAGTGTATCTCGATTTTGGGAATCTTAGCCATGTCGGTAAATATGACATATC 3654
Db 3620 ACTTTAGGCGCGCTAGATCTAGGCAATCCAAATGATACCTGCAAGATGATGATGTC 3679
Qy 3655 TTTGTCAATGTATGACCCCGTACAGAAACCATCACTTACCAAGTGTGAGATCAAGCT 3714
Db 3680 TTGTAATATTAAGAAACAGTACCGCTACCAACAAGTATGAGAGTGTGAGACGACGCC 3739
Qy 3715 ATCCACAACAGATCTTACGTTAAGCTGTGCAACCACTGAAACAATGCGGAAATGT 3774
Db 3740 ATTAAATGAGATGTTGACCAAGAAAGGCTGCTGCACTGAAACCCCGAGGAACTGT 3799
Qy 3775 GTGCTATTAAGGTATGCGCTTGTGATTCGCGCAACCGAAATATCATCTGCGGTGGCA 3834
Db 3800 GTAGCAATTTGTTACGCTATGTGCGAACCGGGCAAGTAAAGCAATCAATGATGCAATGCT 3859
Qy 3835 CGCTATTTAGTTTACCCGCTGTCTGACGCTTAAAGAACTGCCGAAATTAATGAGTT 3894
Db 3860 CCGGAGTTCAAGTTCTGAGGATATGCAACCGAAGTGTCTTAAGAGAGAGACGAAAGTG 3919
Qy 3895 CTCTTCGTTTCTTCCGAGAGCAACGCGAACCAACACATATCAAGACAGACTCGGT 3954
Db 3920 CTATTTGTCTTCAATGGGTGATGATCGTAAACGCAACCATTAACCATTAAGCTCTCC 3979

[illegible][illegible]

Dh 156 CCGTCGTTGAGTGGAGTCATTCAGAGTCAACCAATGACATGCAATGCCAGACA 215
Qy 151 TTTTTCGATGTGGCAACAAAGCTCATTTGAGGCCAAGTCGCCGGACCAAGTTATCTTG 210
Db 216 TTTTTCGACCTGGCTACCAATTTGATCGAGAGAGACTGCAAAAGACACTATCTTG 275
Qy 211 GACATTTGAAAGTGGCCCGTGCAGATGCACTTTCATTCACCGCTCATCTTGTATCTGC 270
Db 276 GATTCGGAGTGGCCCTTCAGAGAAATGATGTCTACGCCAATTAACCACTGGGTATGC 335
Qy 271 CTATGTAAAGCGCTGAAGACCCGGACAGACTTACACGGTATGCGAAGAAGCT-----T 324
Db 336 CCTATGGCAGCGCAGAAAGACCCGAAAGGCTCATAGCTTACGAAAGAACTGGCAGCG 395
Qy 325 AAGAAAGTGCATTTACCGAATAAGACATAGCTCTTAAGGGGAGACCTGTGGAATGC 384
Db 396 GCCTCCGGAAAGGTGGATGAGAGATCGAAGAAATAATCACCGACCTGCAACCGTC 455
Qy 385 ATGTCAACACGAGACGAGAGACTCATCTCTGTGTATGACACAGACGCGCTGTAGG 444
Db 456 ATGGCTACGCCAGACGCTGAATCTCTTACCTTTGGCTGCACTACAGAGCTACGTTGCT 515
Qy 445 TACTTTGAAAGTGTAGCATATACAGAGATGTATACGAGTCCATGCAACGACATCAATC 504
Db 516 ACGGCAACCGAAAGTGGCGGTATACAGGACGTGTATGCTGTACATGCAACAACTCGCTG 575
Qy 505 TACCACACAGGCGCTTAAAGAGTTAGAGCAATTTACTGGAATAGGCTTTGACACACCCCT 564
Db 576 TACCATAGGCGGATGAAGGTGTAGAGACGCGGTATGGAATGGGTTTGAACACACCCCG 635
Qy 565 TTTATGTACAAATAATGATGAGTTCCTTCACTTCAACAAGAAAGTGGCTGACGAG 624
Db 636 TTTATGTGTGACGCGCTTACAGCGCGCTGTATCAACCTTACGCAAAACTGGGCGAGAG 695
Qy 625 AGAGTATTTGAAGCAGCTAACATTGGCCCTCGTAACTCATCTTCAGAGAGCAGGCTT 684
Db 696 CAGGTGTACAGGCGCAGAAACATAGACTGTGTGACAGATCTTGACTGAGGAAAGATC 755
Qy 685 GGAATACTCTCAATCTTGAAGAAAGAGGCTCCACCTACTAATTAAGTATATTTCTCG 744
Db 756 GGGAACTGTCAATCTTCCGAAAGAGCAATTTGAACCTTGCACACAGTATTTCTCG 815
Qy 745 GTTGTGTCAACAATCTTACAGAGATAGATCACTGTTACGTAGCTGGCATCTTCCAAC 804
Db 816 GTAGGATCTTACATTTGTACCTGAGAGCAAAAGCTACTGAGAGACTGGCACTTACCCTCC 875
Qy 805 GTGTTCACATTGAAAGGAAGTCTAATCTCAAGGTAGATGTGGACCATTTGACGCTGT 864
Db 876 GTATTTCACTGAAAGGTAAACATCTTACTGTAGGTGCGATVCAATCGTATCATGT 935
Qy 865 GAAGGATACGTCACTAAAGAAATTAACGATCAGCCCAAGACTTACGGTAAAGTTGAGAAC 924
Db 936 GAAGGATACGTATTAAGAAATCACTATGTGCCCGCTGTACGGTAAACCGTGAAGG 995
Qy 925 TTGCGCTCAACAATGCAATCGGAGGGTTTCTTGAATTGCAAAAGTCAAGATACGCTGGC 984
Db 996 TAGCGCTGAGATCTACGCGGAGGAGATTTCTAGTGTGCAAGACACAGACACTGTCAA 1055
Qy 985 GCGGAGAGGGTTTCTTGTGCTGTGTATCGTATGACCGACCACTTTGGCATCAGATG 1044
Db 1056 GGAAGAAAGATCTATTCCTGTATGCACTACGCTCCCTCAACATCTGTATCAATATG 1115
Qy 1045 ACAGGAGATTTCTGGCACTGACGTTAGTGTGATGACGACAAACATATTGGTGGGCTC 1104
Db 1116 ACTGGCACTATGAGACCGAGCTGACACCGGAGAGCGACAGAAAGTTGTGTGGGATG 1175
Qy 1105 AACCAAGATTTGCTCAATGTAGAGCGCAAGAAATATACTACACAATGACAGAACTAT 1164
Db 1176 AATAGAGAGATTTGTGAACGGAAGAACACAGGAAACCTTAAACGATTAAGACTAT 1235
Qy 1165 CTATTACAGAGTGTGCGCCAGGCGTTTTCAGGTGGGCGCTGAACTGTGTGCGCACTTG 1224
Db 1236 CTGCTTCGATTTGTGGCGGTGGCATTTAGCAAGTGGGGAAGTAATCAAGGACGACCTT 1295

Qy 1225 GAGCAGAGAAAGAACTAAGGGGTGGGAGCGCACTTTACTATAGGCTGTGCTGGCT 1284
Db 1296 GATGATAAAAACCTCTGGGTGTGCGAGAGGTCACTTACTTGTCTGTGTTGGGCA 1355
Qy 1285 TTCAGAGCCAGAAATACATCCATCTACAGAAAGCTGTATCGCAACAAATTAAGAAA 1344
Db 1356 TTTTAAACGAGGAAGATGCAACCATGTACAGAAACCAAGAACCCAGACAAATGTGAAG 1415
Qy 1345 GTACCTGCCCTTTTACTCATTTGTGATTCACGCTTACAGCCAGCGGCTGATATG 1404
Db 1416 GTGCTTTCAAGATTAACTGTGTGTCATCCAGCTTATGTGTACAGGCTTCGCAATC 1475
Qy 1405 GCGTTCCGCGGTAGGCTCAAGCTGTCTTGAACCAACTGTCAACCCGCGCTGATT 1464
Db 1476 CCACTCAGATCAGCATTTAAAGATGCTTTTGGCCAAAGAACCAAGC--GAGATTAAITA 1532
Qy 1465 ACAATGCGCATGTGAGCATCTGCGTGTACAGCAAGAGCTGAAGAGTGGCTGCA 1524
Db 1533 CTGTCTCGACCGGTGTGACGCGAGGATGTGAAACAGAGGAGAGAGAGTGGAG 1592
Qy 1525 GCGGAGAGATCAGAGAGCCCTGCAACCTTGTCTCCCTGAATAGAAAAAGAGCCGTA 1584
Db 1593 GCGGAGCTGACTAGAGAGCTTACACCCCTGCTCCCATCGCGCGGCGAGACGGGA 1652
Qy 1585 GAGGCAAGTAAAGCTTATGCAAGAG-----GCAAGAGCAGGTAGCTGGAGACA 1638
Db 1653 GTCTCGACGCTCGAGCTTGAAGAACTTAGATATCACAGAGTGCAGGGGTGTGGAAACA 1712
Qy 1639 CCAAGAGACACATCAGGGTGACAAAGTTACCCAGGCGAAGAGATTTGGGTCTTACGCT 1698
Db 1713 CTTGCGAGCGGTTGAAGATGTCACCGCACCGCAAGCAAGTACTACAGAAATTAAGTA 1772
Qy 1699 ATACTTCACTCCAGGCGGTATGATATGTAAGAAAACTGCGGTATATCCACCTTGGCG 1758
Db 1773 GTTCTGTCCCGAGAGCTGTGTCAAGAGCTCCAAAGTTGCCCCCGTGCAACCTCTGCA 1832
Qy 1759 GAACAACTACTGGTAATGATCTCAAAAGGTAGGGCAGGAGATCAAAAGTGCAGCTAC 1818
Db 1833 GAGCAGGTGAATAATTAACATTAACATTAACGAGAGGCGCGGCTTACAGGTGCAGGATAT 1882
Qy 1819 CACGTAAGTCAATTTGACAGAAAGGACGCGGTCCCTTTCAGAACTTCCAGGCAATG 1878
Db 1893 GAGGCGAGGGTCTTCACTACATGATGATGATGCGCATTCGCGTCCGTATTAAGCTTTG 1952
Qy 1879 AGTGAAGCGCTACGATGTTTTCACAGAGGAGTTCGTAACAGATACCTGCACAC 1938
Db 1953 AGCGAGAGCGCACTATGTGTACAAAGAAAGGAGTTGCTCAACAGAAACTATACAT 2012
Qy 1939 ATGCAATCAACGAGAGGCGTAAACATGACGAGAGATCTAATGACTGTAAAGACT 1998
Db 2013 ATTTGCGTCAACGAGCCGTGCTGTAACCGAGAGAGAACTTACGAAAGTCAAGACT 2072
Qy 1999 CAGGACACAGACTCAGAAATAGTCTTTCGATTTGACGCAAGAAAGTGTGTAAGCGAGAA 2058
Db 2073 GAAAGAACTACGCGGAGTACGTGTTCGATGATGATTAAGAAAGTGTGCGTCAAGAGAG 2132
Qy 2059 GAGCGAGGTCCCTTGTGCTTAAACCGGTATCTGTGATATCAACATTTCAAGATTGGC 2118
Db 2133 GAAGCGTGGGTTTGGTGTGTGAGAGACTTAACCAACCCCGTGTCAATGATGGCC 2192
Qy 2119 TAGGAGTCTCAAGACAGACGACGACCTCAACAAATCCCAACATCGAGGTAT 2178
Db 2193 TACGAGGGCTGAAGATCAGGCGGTGGCACTATTAAGACTACAGTATGAGGATCTTT 2252
Qy 2179 GAGTGCAGAGTTCAAGTAAATCTGATATCATCAAAAGCCGTGTGACTAAGAAAGATCTG 2238
Db 2253 GGGGTTCCGGGATTCAGCAAGTCTGTATATTAAAGCCTCGTGAACCAACAGACTG 2312
Qy 2239 GTTGTAGTGCAGAAAGAAAGAACTGCGCAGAAATCATCAGGATGTAAAGAGATGAGA 2298
Db 2313 GTCAACGCGGCAAGAGAGAACTGCGCAGAAATAGTTAAAGACGTGAAGAAAGCACCGC 2372

OY	2229	GGTATGATGATGTCGCTAGACACTGTGCATTCAGTGTCTTAAATGGGGTTAAGCACCC	2358
Db	2373	GGGAAGGGGACAAGTAGGGAAAACAGTACTCCATCCGTAAACGGGGTGTGCTGTGCC	2432
OY	2359	GTTAAACCTCTGTACATGATGAGGCAATTGCGCTGCATGACGGGACGCTGTGGCACTG	2418
Db	2433	GTGGACATCTCTATATGTGTGACGAGGCTTTCGTTGCCATTTCCGGTACTCTGTGGCCCTA	2492
OY	2419	ATTGCAATCGTCAACCT--AAGAAATGGTATTTGTGCGGGGACCCAAAACAATGCGGC	2479
Db	2493	ATTGCTCTGTTTAAACCTCGGAGCAATGGTGTTATGCGGAGACCCAAACAATGCGGA	2552
OY	2476	TTCTTTACATGATGTGCTGTAAGTAATTTTAACTGACTATGTGACTGAAAGTATAC	2535
Db	2553	TTCTTCAATATGATGCACTTAAGGTGAATCTTCAACCAATCTGTGACTGAAAGTATGT	2612
OY	2536	CATAAAAACATCTCTAGGAGGTGCACAGACTGTAACCGCATGTCTCCAGCTCTTC	2595
Db	2613	CATAAAGTATATCTCAGACGTTTGACGCGCTCAGTACGCGCATGTCTTACGTTGCAC	2672
OY	2596	TACGACAAGCGAATGMAAGACGGTTAACCATGTGTGCTGATTAATCATCATGATACACA	2655
Db	2673	TACGAGGCAAGATGCGCACGCAACCCGTGCAACAACCAATATCATGACACACA	2732
OY	2656	GGGACACAAAGCCGCAACAAAGATGATCTGATTCTAACCTGTTTCAAGATGGGTGAAA	2715
Db	2733	GGACAGACCAAGCCCAACCCAGAGACATCGTGTTTAATCATCTTCCGAGCGTGGCAAG	2792
OY	2716	CAGCTACAGATTGACTCAAAAATCAGAAATCATGACTGTGGCTGTCTGCCAAGACTT	2775
Db	2793	CAGCTGCATTGTGACTACCGTGACAGAAGTCAAGACAGACAGCAATCTCAGAGGCTC	2852
OY	2776	ACGCGGAAGGCGTTTATGTCTGTCAAGTACAAAGTCAACGAAATCCACTTACCTGCAG	2835
Db	2853	ACCGCAAAAGGGGTATATACCGGTGAAGGACGAAGGATGAAATCCCTGTATGCCCT	2912
OY	2836	ACTTGTGACGACGTTGACGTGTACTTACACGCACAGAAAAACGATGTCTGAAAGAC	2895
Db	2913	GCCTCGGACGACGTGAATGTACTGTGACGCGCACTGAGATGAGCTGTGTGAAAAAG	2972
OY	2896	CTAGCTGTGATCCCTGTGATTAAGCACTTACAGTAAATATCCGGGATTTCAAGGCT	2955
Db	2973	CTGGCGGCGATCCCTGATTAAGGCTCTTCAAACTTCCACAGGGTAACTTTACGGCC	3032
OY	2956	TCATTGTGACGACTGGACGCGAACAACAGCCATTATGTGACACGGGTTCTGATTAAGCG	3015
Db	3033	ACATTGTGAAGATGACGAAGAAAGAACACACAACAAATATGAAGGTATGAAAGACCGCT	3092
OY	3016	CAGACAGCTGATGTGTTCAGAAATTAAGGTGAACGTCTGCTGGGCGAAGGCTTTAGAGCA	3075
Db	3093	GCCTCTGTGACGCGTTCCAGAAACAAGCGAAGCTGTGTGGGCGAAGAACCTGTGCT	3152
OY	3076	GTCTTTGGCCACGGCCCAATGTGTGCTGACGAGACAGCAGTGTGGAGAC--GTGACACCA	3132
Db	3153	GTCTGTGACACTGCGCGAATCAGATTGACAGAGGAGTGTGAGCACCATTAATACAGCA	3212
OY	3133	TTCAAGCATGACAGCGCTACTCAACCTGAATATNGCACTTCTTTTGACACAGGCTC	3192
Db	3213	TTTAAAGAGGACAGAGCTTACTCTTCAAGTGTGCTTTGAATTTGACACAAGTAC	3272
OY	3193	TTTGAAGTGAACCTGACAGTGGGTTATTTTCCGCTCTTACCGTGCACCTTACTTACAGG	3252
Db	3273	TATGAGATTGACCTGGAAGTGGCGCTGTTTTCTGCCCCGAAGGTGTCCCTGTATTAAGAG	3332
OY	3253	GATCAGCACTGTGGGTAATCTCGCCAGGGGAAGAACTATATGGGCTTAAATAGAGAGTATGA	3312
Db	3333	AACAACCACTGGGATMAACAACCTGTGTGAAGATGTATGATTTCAATGCGCAACAGCT	3392
OY	3313	AAGAGTGTACAGCGCATATCCGTCATCAAAAAGCGTTGACACAGGACGGGTAGCT	3372
Db	3393	GCACAGCTGTGAAGCTAGACATACCTTCTGTAAGGGGACGTGGCATACGGGACAGACGCA	3452
OY	3373	GATATTAAGGAATTAATCATCAAGAGACTACTTCCAAACAATTAATGTGTTCATTAAT	3432

Db	3453	GTATTCGCAAGAAAGAAAAATCCACCGCTTTCTGTCTGTGACCATATTAATTCCTATCAAC	3512
OY	3433	CGCGGTTGCCCCACTGTTGATCGTTGACCAAGAGACAGGGTCAACTGATCAACG	3492
Db	3513	CGACGGCTGCCGACCGCCCTGTGTGGCTGAGTACAAACGGTTAAAGCAGTAGGGTTGAG	3572
OY	3493	GGATTCCTATCTAAGATGAAGGGCAATCTGTGTGGTGAATCGCGCATCCTATCAGACTT	3552
Db	3573	TGGCTGGTCAATAAAGTAAGAGGGTACACAGTCTCGTGGTGAAGTAAACCTGGCT	3632
OY	3553	CCAGGGAAGAAAGTAGAGTCCATGGGTCCATTGCCCATTAATC-----CATCAGG	3603
Db	3633	TTGCTTCGACGCAAGGGTCACTTGGTGTCAACCGCTGAATGTCAACAGCGCCGATAGGTGC	3692
OY	3604	TGTATCTCGATTTTGGGAATACCTAGCGCATGTGGGTAAATATGATTAATCTTTGGCAAT	3663
Db	3693	TACGACTTAAGTTTAGACTCCGGCTGACGCGCGAGTTGCACTTTGGTCTTTGTGAAC	3752
OY	3664	GTTAAGACCCCGTACAGGAACCATCACTAACCAAGTGCAGAGATCAGCTATCCACAC	3723
Db	3753	ATTACACGGAATTCAGAAATCCACACTACACAGATGTGTGACACACGCATGAAGCTG	3812
OY	3724	AGCATGCTTAACGTGTAAAGCTGTTCACACCTGCAACATGCGCGGAACATGTGTGCTATA	3783
Db	3813	CAGATGCTTGGGGAGATGCGCTACGACTGTAAACCCGCGCGGCACTCT---TGATGAGA	3869
OY	3784	GGGATGAGGGCTTGCTGATCGCGGAACCGGAATATCATCATGCGGTGCGACAGCTCATTT	3843
Db	3870	GCTTACGGATACCGCCGATTAATAATACAGGAAGCCGTGTGTTCTCTTAAAGCAAAAGTTC	3929
OY	3844	AGGTTTACCCGTGCTGTCTGACGCTTAAGAAACCTGCGCAAAATACTAGGTTCTCTTCGTG	3903
Db	3930	TGCTCTCAGAAAGTGTGGCCCGGATTTGTGTACACAGCAATACAAAGTGTCTTCTGCG	3989
OY	3904	TTCTTGGGCAAGGCAACCGGCAACCAACATGACACAGACAGACTCGGTGATGCTTT	3963
Db	3990	TTCTCCAACTTTGACAAACGGAAAGAGACCCTTACGCTACACCAATGTAATCCAAAGCTG	4049
OY	3964	GACAAATCTATCAAGGGGTCAACAGGTTCAGGGCGAGGAGAGCTCAACGATACAGAGTG	4023
Db	4050	AGTGCCGTGTATGCGGAGAAAGCCATATGACACCGCCGGGTGTGCACTATCAAGAGTT	4109
OY	4024	ATCAGAGTGATCAATTAGCAAGACGCTGACCAAGCTATGTTAATGCTCTAATAGCAAA	4083
Db	4110	AAGAGACAAACATAGCCACGTGCAACGAAGCGGTGTGTGTTAACGCACTAACCGCCCT	4169
OY	4084	GGTCAACCAAGTTCCGGAGTGTGCGGTGCACTGTACCGAAAAATGCGCGGCTGTTTGTAT	4143
Db	4170	GGAATCTGATAGGGATGCGGTATGCAAGGCGCGTGGCGAAMAAATGCGCGTCAAGCTTTAAG	4229
OY	4144	AGACAGCCATTAAGTGTCCGGAACGCTGAAGCTTTGTGAACAGAACCG--CTCATCATTA	4200
Db	4230	GGAGCAAGCAACACAGTGGGCACAATTAACACAGTATGTGGCGCTGTAAACCCCGTCATC	4289
OY	4201	CATGCTGTAGAGCCCAATTTTCTAAGATGCGCGGAACCGAGGGCGCACTTAAGCTCGCA	4260
Db	4290	CAGGCTGTAGCGCTTAATTTCTCTGCAAGCTGAAGCGGAAGGGGACCGCAATTTGGCC	4349
OY	4261	GCTGCTCAATGAGCATATGCGTCAATGCTCAACGCTGAACGGAATTAACAAAAATATCAGTA	4320
Db	4350	GCTGTCTACCGGGCAGTGGCCGCGGAATTAACAGACTGTCACTGAAGCAGCTAAGCATTC	4409
OY	4321	CCGCTACTGTCAACCGGCACTTAATTTCTGTGTGGAAAGATCGAGTGTGATATCATTTGCAT	4380
Db	4410	CCGCTGTGTGCCACAGAGTGTTCACGCGCGGAAGAGATAGGCTGTGAGCAATCCCTCAAC	4469
OY	4381	CACCTGTTCACTCTTTTCGACATTAAGATGCGAGTGCAGCATTAATATGCTTGGATAA	4440
Db	4470	CATCTATTCAACAGCATGACGCAAGGACGCTGAAGTBAACAATTACTGCAAGAGCAAA	4529
OY	4441	CAATGGAGACCAAGATTAATCGAGGCATTCACCGCAAGAAAGCGGTGAAATTTGTGAT	4500

Db	4530	AGTTGGAGAAAGAAAATCCAGAAAGCATTGACATGAGGACGGCTGTGGAATTGCTCAAT	45589
Qy	4501	GATGACAAAGCCAGTAGACATTGACTTGGTCAAGGGTCCACCCAAACAGCTCTTTTGGCAGGC	45606
Db	4590	GATGACGTGGAGCTGCACCAAGACTTGGTGTAGAGTGCACCCGGCAGCAGCCTGTGTGGGT	46459
Qy	4561	AGACCAAGTTACTCCGTCATGAGGGGCAAGTTGATTATATACCTCGAAAGGTACACGATTC	46520
Db	4660	CGTAAGGCTTACAGTACACATGACGGGTGCGCTGTACTGTACTTTGAAGGTACGAAATTC	47099
Qy	4621	CATCAGACCGGCCAAGACATTGCCGAAATCCATGCAATGTGGGCCCAACAAATCTGAGGT	46880
Db	4710	AACACGCGCTGATTATGATATGGCAGAGATCTGACGTGTGGCCAGACTCAAGAGGCA	47659
Qy	4681	AATGACGAGATTTGGTTGTATACATCTCGGGGAGAGTATGTCAGATCTCGCTCCAAATGC	47480
Db	4770	AACGACACGATATGTCTATATACCGCTGGGGCGAAACAATGACACAAATCATGATCCAAATGT	48229
Qy	4741	CCAGTAGAGAGTCAGAGCGGTCTGGCTGCACCTCAACACACTTCCATGCTGTGTAATTAC	48000
Db	4830	CCGGTGAACGATTCGGATTCATCAACACCTCCACAGACAGTGGCCTGTGCGCGTAC	48889
Qy	4801	GCTATGACGGCTGAGCCGCTATACAGGTGTGGGCTCTGCGAAGAAAAGACAGTTGGCGCTA	48660
Db	4890	GCAATGACAGACGAACGAGATGCCCCGCTTAGGTATCACCAAGTTAAAAAGCATGGTGGTT	49499
Qy	4861	TGCTCATCATTTCTGTTGTCGGAAGTACAGATCAACAGCGCTGAGAAAGCTACAGTGCAGC	49230
Db	4950	TGCTCATCTTTTCCCTCCCGAAATATACATGTAGATGGGGGTGAGAAAGTAAAGTGCAG	50099
Qy	4921	AAACCAAGTCTGTTTTCAGGGCGGTATCACCGGCTGTACACCCACGAAATGACGGGAA	49880
Db	5010	AAGGTTCTCTGTGTGACCCGACGGTATCTTCAGTGGTTAATGTCCCGGGAATATATGCGCA	50659

RECORD 5
 US-08-466-277-1
 ; Sequence 1, Application US/08466277
 ; Patent No. 6190666
 GENERAL INFORMATION:
 APPLICANT: Garoff, Henrik
 Liljestrom, Peter
 TITLE OF INVENTION: DNA Expression Systems Based on
 Alphaviruses
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Birch, Stewart, Kolach & Birch
 STREET: P.O. Box 747
 CITY: Falls Church
 STATE: Virginia
 COUNTRY: USA
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/466,277
 FILING DATE: 06-Jun-1995
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/920,281
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Murphy Jr., Gerald M.
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 828-103P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-241-1300
 TELEFAX: 703-241-2648
 TELEX: 246345
 INFORMATION FOR SEQ ID NO: 1:

```

? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 11517 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: RNA (genomic)
? -
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Semliki Forest Virus
?
? FEATURE:
?
? NAME/KEY: -
? LOCATION: 1..11517
? OTHER INFORMATION: /label= genome
? /note= "Semliki Forest Virus complete nucleotide
? sequence, presented as a cloned DNA sequence; see
? Figure 5."
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: 87..7379
? OTHER INFORMATION: /product= "SFV polypeptide"
?
? FEATURE:
?
? NAME/KEY: CDS
? LOCATION: 7421..11179
? OTHER INFORMATION: /product= "SFV polypeptide"
?
? SEQUENCE DESCRIPTION: SEQ ID NO: 1:
?
? US-08-466-277-1

```

Query Match	31.3%	Score 1564	DB 3	Length 11517
Best Local Similarity	58.3%	Pred. No. 0		
Matches 2004	Conservative	0	Mismatches 2040	Indels 36
			Gaps	8
QY	AGAATTACGTTGACTTAGATGCTGACACCCCGTATGTCAAGTCGTTACAGCGACGTTT	90		
Db	AAAGTGCATGTTGATTATGAGGCTGCAGCCCATTCATCAAGTCTTTGCGAAGAAAGCATTT	155		
QY	CCACAAATTGGATGCAAGCAAGGCAAGTCACTGACAAATGACATGCCAAATGCCAGAGCG	150		
Db	CCGTGTTGAGGTGAGTCAATTCGAGGTCAACCAAAATGACCAATGACAAATGCCAGAGCA	215		
QY	TTTTGCGATGCGCAACAAAGCTCATTTAGAGCGAAATGCAACCGGACCAAGTTATCTTG	210		
Db	TTTTGCGACCTGGCTCAACCAATTGATGAGCAGAGACGTGCAAAAGACACACTCATCTTG	275		
QY	GACATTGGAAGTGGCCCGCTGACAGATGCAACATTCCAATCAACCGCTATCATTTGATCTGC	270		
Db	GATATCGGACAGTGGCCCTTTCGAGGAAATGATGTTTACGCAACAAATACCACTGCGCTATGC	335		
QY	CCTATGATTAAGCGCTGAAGAGCCCGGACAGACTCAACAAAGTTATGCAAGAAAGT-----T	324		
Db	CCTATGCCAGGCGAAGAACCCCGAAAGGCTCGATAGCTACGCAAGAACTGGCAGCG	395		
QY	AAGAAAGTGCATTACCGACAAAGAACTAGCCTTTAAAGGGCGAGACTGCTGGAAGTC	384		
Db	GCCTCGGGAAGGTGCTGATAGAGATGCGACGGAATAATCAACCGACTGCGACACCGTC	455		
QY	ATGTCAACACAGACGCGAGAGATTCCTCTGTGTATGCAACAGACGCCACGTATAGG	444		
Db	ATGGTACGCGACAGAGCTGAATCTCTCACTTTTGCCTGCAATACAGAGCTCACTGGTCTCT	515		
QY	TACTTTGGAAGTGTGACGATATACCAAGATGTGACGCGAGTCCATGACCCGATCAATC	504		
Db	ACGGAGCGCAAGTGGCCGTATACAGGACGTTATGCTGTATCATGACCAACATCGCTG	575		
QY	TACCAACCAAGGCGCTTAAAGAGTTTAGCAATTTACTGGATTAGGCTTTGACACGACCCCT	564		
Db	TACCAATCGCGATGAAGAGTGTCAAGACGCGCTTTGGATTGGGTTTGACACCAACCCCG	635		
QY	TTTATGTCAAAAACATGCGAGGTTCTTACCCCTACTTTCAACAACGAACTGGGCTACGAG	624		
Db	TTTATGTTTGGACGCGCTAGCAGGCGCGTATCCAACCTACGCCCAAACTGGGCGACGAG	695		
QY	AGAGTATTGGAAGCAGTACATGTCCTCGGTAACTCAGATCTTCAGAGAGCAGGCTT	684		

Db 696 CAGGTTACAGGCGAGGAAACATAGACCTGTGTGACGATCCCTTGAAGTGAAGGAAAGACTC 755
Qy 685 GGAATACCTCAATCTCTTAGAAGAGAGGCTCAACCTACTAATATAGATCATATTTCTCG 744
Db 756 GGCATACTGTCTCTCGGAAAGAACATTGAACCTTGCAGACAGCTCATGTTCTCG 815
Qy 745 GTTGTTCAAACATCTACACAGAAAGATAGATCACTGTTACGTAGCTGATCTTCCAAAC 804
Db 816 GTAGATCTACATTTGTACCTAGAGCAAGAAAGCTACTAGAGAGGCTGAGCACTTACCTCC 875
Qy 805 GTGTTCACCTTGAAGAAAGTCTAATTCAAGTAGATGTGGGACCATTTCTCACTCT 864
Db 876 GTATTCCACTGAAAGGTAAACATCTTTACCTGTAGGTGCAATACCTATCATGT 935
Qy 865 GAAGGTTGCTCATCAAAAAGATTAAGATCAAGCCAGGACTATAGGTTAAATTGAGAAC 924
Db 936 GAAGGTTAGTATGTTAAGAAATCACTATGTCCCGGCTGTAGCGTAAACGGTAAAG 995
Qy 925 TTGGGTCACATCATCGAGAGGTTTCTTGAAGTGAAGTCAAGATACGCTGCGC 984
Db 996 TACGCGTAGACCTATCAAGCGGAGGAGTTCTTAGTGTGAGAACACAGACACTGTCAA 1055
Qy 985 GCGGAGAGGTTTCTTTTGTGTGTAGTATGTAACAGCCACATTTGCGATCAGATG 1044
Db 1056 GGAGAAAGATCTATCCCTGTATGACCTACGTCCTCCCTCAACCATCTGTATCAATG 1115
Qy 1045 ACAGGATTTCTGCACTGACCTTAGTGTGATGATGACCAAAAATATTTGGGCTC 1104
Db 1116 ACTGCAATCTAGCCAGCACTCAACCGAGAGCGCACAGAAATTTGTATGTGGATG 1175
Qy 1105 AACCAAGGATTTGTCTCAATGTAGAGACGCAAAAGAAATCTAACAATGACAACTAT 1164
Db 1176 AATCGAAGATAGTTGTGAACGAGAAACACAGCGAAACATTAACATGAAAGACTAT 1235
Qy 1165 CTATTACAGTGTGTGCGCCAGGCGTTTTCAGGTGAGCGCGTGAACATGTGCGCAT 1224
Db 1236 CTGCTTCCGATTTGTGGCGTGTGCAATTTAGCAAGTGGGCGAGGAAATACAAAGCACA 1295
Qy 1225 GACGCGAAGAAAGACTAGGGGTGGGAGGCGGACATTTACTATAGGGCTGTCTGGCT 1284
Db 1286 GATGTGAAACCTCTGTGGGTGTGAGAGAGGTACTTACTTGTGTCTGTGTGGCA 1355
Qy 1285 TTCAAGACCCAGAAATCATCATCTACAGAAAGCTGTGATCGCAACATTAAGAA 1344
Db 1356 TTTAAAGAGAGAAATGACACCATGTACAAAGAAACAGACACCCAGACATATGTAAG 1415
Qy 1345 GTACCTGCGTTTGTACTATTTGTGATTCACAGCCTTACAGCCAGCGGCTCGATATG 1404
Db 1416 GTGCTTCAAGATTTAATCTCGTTGTGATCCGAGCCTATGTTACAGGCTGTGCANTC 1475
Qy 1405 GGCCTCCGCGTGTGAGCTCAAGTGTGTGTTGAACCAAGTGTAAACCCGACCGCTATT 1464
Db 1476 CCAAGTCAATCAAGCATTAAGATGCTTTTGGCCAAAGAAAGACCAAC---GAGATTAATA 1532
Qy 1465 ACAATGCGCATGTGTGAGCATCTGTGTGCTTACAGCAAGAAAGTGAAGAGTGTCTCA 1524
Db 1533 CCTGTCTGACGCTGTGTGACGCGAGGATGTGTGAACAAAGAGAGAGAGTGTGAAG 1592
Qy 1525 GCGGAGAGATCAAGAAAGCCCTGACCTTGTCTCCCTGAATAGAAAAGAGACCTTA 1584
Db 1593 GCCGAGTGAATAGAGAGCTTACACCCCTGTGCTCCCATGCGCGCGGCGAGAGCGGGA 1652
Qy 1585 GAGGCAAGATGAGCTCATTTATGCAAG-----GACAGACAGGTAGGTGAGAC 1638
Db 1653 GTCTGTGAGTGTGAGTGAAGAACTAGATATCAAGCAGGTGAGGGGTGTGTGAAC 1712
Qy 1639 CCACGAGACATCAGGAGTGAAGATTACCGAGCGAAGAGAAATGTTGGGTCTTAAAGCT 1698
Db 1713 CCTGCGAGCGGTTGAAGTCAACCGCACAGCGGAACGACGTAATCTAGGAAATTTACGTA 1772
Qy 1699 ATACTTCAACCCGAGCGGTATTGATAGTGAAGAAATGCGGTGTATCAACCATTTGCG 1758

Db 1773 GTTGTGTCCCGACAGCCGTGCTCAAGAGCTTCAAGTTGACCCCGGTGACCTCTAGCA 1832
Qy 1759 GAACAAGTACTGTAAATGACTCAACAAGGTAGGGGAGGAGATGAAGATGAGCATAC 1816
Db 1833 GAGCAGGTGAATATATACATTAACGAGAGGCGGCGGTTTACAGGTCAAGCAAT 1892
Qy 1819 CACGTAAGTCAATTTGATACAGAAAGGAGCGGCGTCCCTGTTCAAGACTTCCAGCAT 1878
Db 1893 GACGCGAGGTGCTACTACATGTGTGATCGGCATTTCCGGTCCCTGAGTTCAACCTTTG 1952
Qy 1879 AGTGAAGCGCTACATGTGTTTCAAGAGAGGAGTTGTGAACAGATACCTGACAC 1938
Db 1953 AGCGAGACCGCCTATGTGTGATCAACGAAAGGGAGTTGTCTCAACGAGAACTAATACAT 2012
Qy 1939 ATCGCAATCAACGAGAGCGCTAAACATGACGAGAGTCTATTAAGCTGAAGCT 1998
Db 2013 ATTGCGTTCAAGGACCGTCCGTGAAACCGACGAGAGAACTACAGAAATGTCAGAGT 2072
Qy 1999 CAGACACAGACTCAGAAATAGCTTTCATATTTGACGACGAAAGTGTGTTAAGAGAA 2058
Db 2073 GAAGAATGACCGCGAGTACTGTTCGATGATTAATAAATGCTGCTCAAGAGAG 2132
Qy 2059 GACCGAGTCCCTGTGTGCTTAAACCGGTATCTGTAGATCCATTTACGAGTTTGG 2118
Db 2133 GAAGGTGCGGTTGTGTGTGTGAGAGCTTAACCAACCCCGTTCATGATTTGCGC 2192
Qy 2119 TACGAGATCTCAAGACACAGACAGACACTTCAAAAGTCCCAACCATGGAATCT 2178
Db 2193 TACGAAAGGCTGAAGATCAGGCGGTGACACATTAAGACTACATGATGAGTCTTT 2252
Qy 2179 GAGGTGCGAGTTCAAGTAAATCTGGAATCTCAAAAAGCGGTGATCTAAGAAAGATG 2238
Db 2253 GGGGTTCCGGATCAGGCAAGTCTGTATTTAAGACCTCTGTACCAACACATCTG 2312
Qy 2229 GTTGTAGTGCAGAAAGAAACCTGCGCAGAAATCATCAGGATGTAAAGAGATGAGA 2298
Db 2313 GTACACAGCGCAAGAAAGAAACCTGCCAGAAATATGTTAAGATGAAGAACCGC 2372
Qy 2299 CGTATGATGTTGCTGTAGAGACTGTGATTCAGTGTCTTCAATATGGGTTAAGACCC 2358
Db 2373 GGGAAAGGAGACAGTGAAGGAAACAGTGCATCTCCTGATTAACGGGTGTGTGTGCTC 2432
Qy 2359 GTTAAACTGTATCATGTATGATGAGCTTCTGCGCATGACAGGAGCGCTGTGACATG 2418
Db 2433 GTGACATCTCTATATGTGAGAGAGGCTTTCGTTGCCATTCGATCTGTGTGCCCTTA 2492
Qy 2419 ATTGCATGTCTAAACCT--AAGAAAGTGTATTTGTGCGGAGACCCAAACATGCGC 2475
Db 2493 ATTGCTTTGTTAAACCTCGAGCAAGTGTGTATGCGAAGCCCAAGCAATGCGGA 2552
Qy 2476 TTTCTTAAATGATGTGCTGAAAGTACATTTTAAACATGACATATGCACTGAAAGTATC 2535
Db 2553 TTTCTCAATATGATGACAGCTTAAGGTGAATTTCAACCAACATCTGCACTGAAGTATGT 2612
Qy 2556 CATAAAGCATCTCTAAGAGGTGACACAGACTGAACCGGATGCTGTCCACGCTCTTC 2595
Db 2613 CATTAAGTATATCAAGCTTGAACCGCTCAAGTACGCGCACTGTCTTACCTTGAC 2672
Qy 2596 TACGACAAAGCAATGAAGCGTTTAAACCATGTGTCTGATTAATCATATGATATACCA 2655
Db 2673 TACGAGAGCAAGATGCGGACAGCAACCGTGAACAAACCATATATATGACACCA 2732
Qy 2656 GGGACACAAAGCCGACAAAGATGATTTGATTTTAACTGTTTCAAGAGATGGTGA 2715
Db 2733 GGAAGAGCAAGCCCAAGCCAGAGACATCTGTATTAATCTTCCGAGGCTGTGGCAAG 2792
Qy 2716 CAGCTTACAGATTGACTCAAAAATCAAGAAATCAATGACTGGGCTGATCGCAAGAACT 2775
Db 2793 CAGCTGAGTGTGACTACCGTGAACAGAAATCAATGACAGAGACATCTCAAGGCTTC 2852
Qy 2776 ACGGGAAGAGCGTTTATGTGTGATCAGGTCAAAAGTCAACGAAATCACTACTGCGAG 2835
Db 2853 ACCGCAAGAGGATATAGCGCGTAAAGGAGATGATGAATATCCCTGTATGCCCCCT 2912

QY	2836	ACTTCTGAGCAGCTGAACTGCTGTACTTTCACGACACAGAAAACGATGCTCGAAGCG	2859
Db	2913	GGCTGGAGCAAGTAATGTACTGCTGACGGCAGCTGAGTAAAGCTGTGTGAAAACG	2972
QY	2896	CTAGCTGTGATCCCTGGATTAAGACACTTACAGCTAAATATCCCGGGGATTTCACGCT	2955
Db	2973	CTGGCCGGCGATCCCTGGATTAAAGGTCTTATCAAAACATTTCCACAGGGTAACTTTAACGGCC	3032
QY	2956	TCATTGGACGACTGGCAGCGGAAACACGACCGCATTTATGGACCGGCTCTTGATTAAGCG	3015
Db	3033	ACATTGGAAAGATGGCAAGAAAGAACACCAAAATTAAGAGGTGATTGAAGACCGGCT	3092
QY	3016	CAGACAGCTGATGTGTTCAGAAATTAAGGTGAACGCTCTGGCGCAAGGCTTTTAGAGCCA	3075
Db	3093	GGCCTGTGGACCGGCTTCCAAACAAAGCGAAGCTGTGTTGGCGAAAAGCTGTGCTCT	3152
QY	3076	GTCCTTGCCACGGCCCAACATTGTGTGTCAGAGACAGCAGTGGAGAC---GTTCACACCA	3132
Db	3153	GTCTCGACACTGCGCGGAATCAGATTGACACAGAGGAGTGGAGACCATATATTACAGCA	3212
QY	3133	TTCAAGCATGACAGAGGCTACTCACCTGAAATNGCACAGTAATCTTTTGGACACAGGTTG	3192
Db	3213	TTTAAAGGAGACAGAGCTTACTCTTCATGTGGGCTTGAAATGAATTTGACCAAGTAC	3272
QY	3193	TTTGGAGTAGACCTGGACAGTGGGTTATTTTCCGCTCTTACCGCTGCACTTACTTACAG	3252
Db	3273	TATGAGATTGACCTGGACAGTGGGCTGTGTTTCGCCCCGAAAGGTGTCCCTGTATTAGAG	3332
QY	3253	GATCAGCACTGGGATTAATCTCCGACAGGGAAGAACATGTATGGGCTTAATAGAGAGTAGCA	3312
Db	3333	AACAAACCACTGGGATTAACAGACTGTGGTGAAGAGTATGGAATTCAAATCCGCAACAGCT	3392
QY	3313	AAGAGTTGTACACGGCGATATCCGTGCATCCAAAGGGGTTGACACAGCAGGGTACT	3372
Db	3393	GCCAGGCTGGAAAGCTAGACATACTTCTGTAAGGGGCACTGGCATACGGGCAAGCAGCA	3452
QY	3373	GATATTAAGGAATTAATACATCAAGACTACTCTCCAAACAATTAATGTGTTCCATTAAAT	3432
Db	3453	GTTATCCGAGAAAGAAAATCCAACCGGTTTCTGTGCTGGACAATGTAAATCTTATCAAC	3512
QY	3433	GGCGGGTTGGCCCACTCGTTGATGTTGTAACCAAAAGGACAGGGTAAACAATGATCACAGC	3492
Db	3513	CGAGGGCTGGCAGCGCCCTGGGTGCTAGTACAAAGACGTTAAAGCAGTAGGGTTGAG	3572
QY	3493	GGATTCTCATCTAAAGATGAAGGGCAAACTGTGTGTGTGATCGCGGATCTTATCAGACTT	3552
Db	3573	TGGCTGTGCATTAAGTAAAGGGGTACACAGCGTCTGTGGTAGTAACTTACACTGGCT	3632
QY	3553	CCAGGGAAAGAAATGAGAGTCCATGGGTCATTTGCCACTAATAC-----CATCAGG	3603
Db	3633	TTTGCTCGACGCAAGGTCATTTGTGTGTCAACCGCTGAATGTCAAGGCGCCGATATAGTGC	3692
QY	3604	TGTGATCTCGATTTTGGGAATACCTTACCATGTGCGGTAAATATGACATTAATCTTTGTCAAT	3663
Db	3683	TACGACTTAAGTTTAGACTGCGGCTACGCGGACAGGTTTGACTTGTGTGTGAAC	3752
QY	3664	GTTAAGACCCCGGTACAGGAACCATCACTTACCAACAGTGCAGAGATCACGCTATCCACAC	3723
Db	3753	ATTCCACAGGAATTCAGAAATCCACACTACACAGCTGTGTGTCAACACGCGCATGAAGCTG	3812
QY	3724	AGCATGCTTAACGTGTGAAGGCTGTCCACACCTGAAACATGTGGCGGAACATGTGTGCTATA	3783
Db	3813	CAGATGCTTGGGGGAGTGGCGCTACGACTGCTAAACCAGGCGGATCT---TGAATGGA	3865
QY	3784	GGGTATGGGCTGTGCTGATCGGCAACGAGAAATATCATCTGCGGTTGGACAGCTCATTT	3843
Db	3870	GCTTACGGATACGCCGATTAATAATCAGGAAGCCGTGTTTCTCTTAAAGCAAGAAAGTTG	3929
QY	3844	AGGTTTACCCTGTCTGTACGCTTAAGAACCTGCCGAAAATACTGAGGTTCTTTCGTG	3903
Db	3930	TGCTGTGCAAGAGTGTTCGCCCGGATTTGTGTCAACACCAATATACGAAGTGTCTTGTCTG	3989

QY	3904	TTCTTGGCAAGGCAACGGCAACACACATGACCGAGCAACATCGGTAGTCTT	3963
Db	3990	TTCTTCAACCTTGGCAACGGAAGAGACCTCTACGCTACACAGATGAATACCAAGT	4049
QY	3964	GACAAACATCTATCAAGGGTCAACCGAGATACGAGCGAGGAGCTCCAGCTACAGATG	4023
Db	4050	AGTCCGTGTATGCCGAGGAAGCCATGAGACAGGCCGGGTGTGCAACCATCTTACAGATT	4109
QY	4024	ATCAGAGGTGACATTAGCAGAGGCTGACCAAGCTATCGTTATGCTGCTATATGCAAA	4083
Db	4110	AAGAGAGCAGCATAGCCACCTGACACAGAACCGCTGTGGTTAACGACGCTTAAGCCCTG	4169
QY	4084	GGTCAACCAAGTTCCGAGTGGGGTGCACTGATCCGAAAAATGGCCGCGCTTTTGAT	4143
Db	4170	GGAACTGAGGGAGATGGCGTATGCAAGGGCGCTGCGCAAGAAATGGCCGTACGCTTTAAG	4229
QY	4144	AGACAGCCAAATAGCTGTGCGACCGCTAGACTTGTGAAAGCAGAACCG--CTCATATA	4200
Db	4230	GGAGCAGCAACACAGTGGGCACAATTAAACAATATGTGCGCTGTACCCCGTATC	4289
QY	4201	CATGCTGAGACCCCAATTTTTCTTAAGATGCGGAAACCGAGGGCGACCTTAAGCTGCA	4260
Db	4290	CACGCTGAGGCGCTTAATTTCTGCGCACGCTAAGCGGAAAGGAGACCGGAATTGGCC	4349
QY	4261	GCTGCTATCATGAGATGACGCTCATCGTCAACGCTGACGGATTACAAAAATATCAGTA	4320
Db	4350	GCTGTCTACGGGCAGTGGCCCGCAAGTAAACAGACTGTACTGACAGCGTACGCTATC	4409
QY	4321	CCGCTACTGTCAACCGGCACTTATTTCTGGTGGCAAGATCGATGATCAATCATTCAT	4380
Db	4410	CCGCTGCTGTCAACAGGAGTGTTCAGCGCGCGGAAGATAGCTGACAGCAATCCCTCAAC	4469
QY	4381	CACCTGTTCACTGCTTTGCACTACGATGCGATGCGCATGCAACATATTTGCTTGATTA	4440
Db	4470	CATCTATTCAAGACATGAGACGCCACGACGCTACGTCACATCTACTGAGAGACAAA	4529
QY	4441	CAATGGAGACACAGATATATGAGGCGCATTCACGCAAAAGAAACGTGCAATTTGCGAT	4500
Db	4530	AGTTGGAGAAAGAAATTCAGAAAGCCATTGACATGAGACGCGTGTGAGTTCTCAAT	4589
QY	4501	GATGACAAAGCAGTAGACATTGACTTGGTCCAGGTCACCCMAACGCTTTTGGCAGCC	4560
Db	4590	GATGACGTGAGCTGACACACAGACTTGGTGAAGTGCACCGGACACAGCCTGTGGGT	4649
QY	4561	AGACCAAGTTACTCCGTCAATGAGGCGAAGTTGATTCATCCGGAAGGTACACGATTC	4620
Db	4650	CGTAAGGCTACAGTACCACTGACGGGTGCGTACTGTGATCTTTGAAGGTACCAAAATTC	4709
QY	4621	CATCAGACCGCAAGACATTGCGCAAAATCATGCAATGTGGCCCAACAAATCTGAGGCT	4680
Db	4710	AACCAAGCTGCTATTGATATGGCAGAGATATCTGACGTTGTGGCCAGACTGCCAAGAGCA	4769
QY	4681	AATGACGAAATTTGCTGTATACCTCTGGGGAGATATGTCAGACATTCGCTCAAAATGC	4740
Db	4770	AACGAACAGATATGCTATACGCGCTGGGCAAAACAAATGCAACAATACAGATCCAAATGT	4829
QY	4741	CCAGTGAAGAGTCAAGAGCGCTGTCTCACTCAACAACCTTCATGCTGTGTAATTAC	4800
Db	4830	CCGGGAACGATTCGATTATCAACAACCTCCACAGACAGTGGCCCTGCTGTGGCGCTAC	4889
QY	4801	GCTATGACGGCTGAGCGGTATACAGGTTGGCTCTGCGAAGAAAGAACAGTTGCGCGTA	4860
Db	4890	GCAATGACAGAGAACGATGCGCCGCTTATGATCAACCAAGTTAAAGCATGTGTGTT	4949
QY	4861	TGCTCATCTTCTGTGTCGCAAGTACAGATCAACAGCGCTGCAAGAGCTACAGTGACG	4920
Db	4950	TGCTCATCTTTCCTCCCGAAATATACATGATGAGGGTGTGCAAGAGTAAATGTTCGAG	5009
QY	4921	AAACCAAGCTGTTTTCAGGCGTGTACACCGGCTGTACACCCAGGAATATAGCGGAA	4980
Db	5010	AAAGTTCTCTGTTGACCCGACGAGTCTTCAATGTTAAGTCCGCGGAAGTATGCGCA	5069


```

RESULT 6
US-09-554-337-1
; Sequence 1, Application US/09554337
; Patent No. 6475780
; GENERAL INFORMATION:
; APPLICANT: Parrington, Mark
; APPLICANT: Li, Xiaomao
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
; FILE REFERENCE: 1038-1042 MS
; CURRENT APPLICATION NUMBER: US/09/554,337
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/065,791
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/CA98/01064
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 15538
; TYPE: DNA
; ORGANISM: respiratory syncytial virus
US-09-554-337-1

```

Query Match	31.3%;	Score 1564;	DB 4;	Length 15538;
Best Local Similarity	58.3%;	Pred. No. 0;		
Matches 2904;	Conservative	0;	Mismatches 2040;	Indels 36; Gaps 8;

OY	31	AGAAATCAGCTTGACTTGTAGTCTCAGACGCCGTATGCAAGTGGTTACAGCGACGTTT	90
Db	2642	AAAGGCATGTTGATATTGAGGCTACAGCCCATTCATCAAGTCTTTTGAGAAAGCAATTT	2701
OY	91	CCACAATTGAGATGAAGCAAGCAGGTCACTGACAAATGACCATGCCAATGCCAGACG	150-
Db	2702	CCGTCGTTCCAGGTGAGATGCATTTGACAGGTCAACCAAAATGACCATGCAATGCCAGACA	276
OY	151	TTTTCCGCAATGGCAACAAGCTCATTTAGACCGAAGTCGACCCGGACCAAGTTATCTTG	210
Db	2762	TTTTTCGACCTGGCTACCAAAATTGATCGACGAGGACTGAAAGAACACACTCATCTTG	2821
OY	211	GACATTGGAAGTCGCGCCGTCGACGATGCACATTCCAATTCACCGCTATCATTTGATTCG	270
Db	2822	GATATCGCAGTGGGCTTTCCAGAGAAATGATGTCTTAGCAAAATACCACTGGCTATGC	288
OY	271	CCTATGATAAAGCGCTGAAGACCCCGACAGACTCAACGGTATGCGAAAGACT-----T	324
Db	2882	CCTATGCGCAGCGCAAGAACCCCCMAAGGCTCGATACTACGCAAAAGAACTGGCAGCG	2941
OY	325	AAGAAAATGACATTAACCGACACAAGAACTAGGCTTAAGGGGGGAGACCTGTGGAAATC	384
Db	2942	GCTTCGCGGAAGGTCTGATAGAGATTCGAGAGAAAATACCGACCTTGACACCGTC	3001
OY	385	ATGTCACACGACGACGACGAGACCTCATCTCTGTGTATGCACACAGACGCCCTGTAGG	444
Db	3002	ATGGCTAAGCAGACGCTGAATCTCTCTCACTTTTCCGCAATACAGACGTACGTCGT	306
OY	445	TACTTTGGAATGTAGCGATATACCAATATGTATGCGAGTCAGACCCGACATCATC	504
Db	3062	ACGGCAGCCGAAGTGGCCGTATACCGAGACGTTATGCTGTACATGCAACCAATCGCTG	3121
OY	505	TACCAACGAGCGCTTAAAGAGTTTGAACAATTTACTGATAGGCTTTGACACGACCCCT	564
Db	3122	TACCATCAGGCGATGAAGGTGTCAAGACGGCGTTATGGATTGGGTTTGACACACCCCG	3181
OY	565	TTTTATGTACAAAAACATGCGAGGTTCTTACCTTACTTACACAGAACTGGGCTGACGAG	624
Db	3182	TTTATGTTTGAACGCGCTAGCAGGCGCGTATCCAACTTCAGCCAACTGGGGCCACAGAG	3241
OY	625	AGAGTATTGGAAGCACTATACATTTGGCTCGGTATCTCAGATCTTCAGAGAGCAGGCTT	684
Db	3242	CAGGTCTTACAGCCAGAGACATAGGATCTGTGTGACAGATCCTTTATCTGAGGAAAGATTC	3301
OY	685	GGAAAACCTCTCAATCTTAAAGAGAAGGCTCCAACTTACTAATTAAGATCATATTCCTG	744

Db	3302	GGCAAACTGTCATTTCTCCGAAAGCAATTGAAACCTTCCGACACAGTATGTTCTCG	3367
Qy	745	GTGGGTTCAACATCTACACAGAAGATAGATCACTGTTACGTAGCTGCATCTTCCAAC	804
Db	3362	GTAAGATCTACATTGTACACTGAGGCAAGAAAGCTACAGAGAGCTGSCACTTACCTCC	3422
Qy	805	GTTGTTCCACTTGAAGAAAGCTTAATCTTACACAGGTAGATGTGGACCATTTGCAGCTGT	864
Db	3422	GTAATTCACCTGAAAGGTAAACAATCCCTTTTACCTGTAGTCCGATACCATGTAATCATGT	3483
Qy	865	GAAAGGTAACGTCATCAAAAAGATAACATACGCCAGGACATATACGTAAGCTTGAAAC	924
Db	3482	GAAAGGTAACGTAAGTAAGAAATACTATGTGCCCCCGGCTGTACGGTAACCGTAAAGG	3543
Qy	925	TTGGCCGTCACAAATGCAATCCGAGGGGTTTCTTGAGTTGCAAAAGTCACAGATACGCTGCG	984
Db	3542	TACCCCGTACGTAACACGGGAGGAAATCTTAATGTGCAAGACCAACAGACATGTCAA	3603
Qy	985	GGCAGAGGGTTCTTTTGTCTGTGTATCGTATGTACACGCACACTTTCGATCAGATG	1044
Db	3602	GGAAAGAAGCTCATTTCCCTGTATGCACTACCTCCCTCAACCATCTGTGATCAATG	3663
Qy	1045	ACAGGGAATTCGGCACTGACGTTAAGTGTGAATACCGACAAAACATAATGGTGGGCT	1105
Db	3662	ACTGCACTACGACCGACCGTACACCGGAGACGACGAAGTGTATGAGTGGAATG	3723
Qy	1105	AACCAAAAGATTTGTCATATGTATGAGACGCAAGAAATACTAACACATGACAGACTAT	1164
Db	3722	AATCAGAGATAGTGTGTGACGGAAGAACACAGGAAACCTAACACATTAAGAACTAT	3783
Qy	1165	CTATTTACCAATGTCGCGCCAGGCGTTTTCACAGTGGGCGCGTGAAATCTGCGCATG	1224
Db	3782	CTGCTCCGATTTGTCGCGCGTATTTAGCAATGGGGAAGGAATACAAAGCAGACTT	3843
Qy	1225	GACGACGAAGAAACTTAGGGGTGCGGAGCGCACTTTACTATGGGTGCTGCTGGCT	1284
Db	3842	GATATGAAAAACCTCGGGGTGCCGAGAGGTCACTTACTTTCGTGCTTGTGGCA	3903
Qy	1285	TTCAAGACCCGAAGAAATCATCATCATCTACAAAGACCTGTGACGAAACATAATGA	1344
Db	3902	TTTAAACGAGAAAGATGACACATCATGTACAGAAACCAAGACACCACTAATGTGAG	3963
Qy	1345	GTACCTGCGCTCTTGACTCATTTTGTGATTCACAGCCTTACACGCCAGCGGCTCGATG	1404
Db	3962	GTGCTTCAGAGTTTACTCGTGTGTCATCCGAGCCTATGCTACAGGCGCTCGCAATC	4023
Qy	1405	GGCTTCCGCGTAGGCTCAGCTCTGCTGTGAACCAACTGTCAAACCGCACCGCTATT	1464
Db	4022	CCAGTCAGATCACCATTTAAGATCTTTTGGCCAAAGAACCAAG---GAGAGTTAATA	4083
Qy	1465	ACAAATGGCCGATGTGAGACATCTGCGTGGCTTAAGACGAAGACGTAAGATGGCTCA	1524
Db	4079	CTGTTTCTCGACGCGTGTCAAGCCAGGAGTCTGAACAAGGAGAAAGAAAGATTGAG	4139
Qy	1525	GCGGAAGATCAGAGAAAGCCCTGCGCAACCTTGTCTCTGTAATAAGAAAAAGAACGTA	1584
Db	4139	GCGAGCGTACTAAGAAAGCTTACCAACCCCTGTCCTCCCATCGGCGCGGAGAACGGGA	4199
Qy	1585	GAGCAGAGTAGACCTCATTTATCAAGAG-----GCAGAGCAGGTAGCGTGAAGCA	1638
Db	4199	GTCCTCAACGTCGACGTTGAAGAACTGAGATATCACACAGGTGACAGGGGTGTGAAACA	4258
Qy	1639	CCAGAGAGACATCAGAGGTGAACAAGTACCAGCGGAAGAGAAATTTGGTCTTAAGCT	1699
Db	4259	CTTGCAACCGGCTTGAAGATCACCGCACACCGCAACGATCTATAGGAATAATTAAGTA	4319
Qy	1699	ATACTTTACCCCAAGCGGATTTGAATAGTGAATAAATGCGGCTGATATCCCAATGCGC	1758
Db	4319	GTTCTGTGTCGCGAGACGCTGCTAAAGAGTCCAAAGTTTGGCCCCCGTCAACCTCTAGCA	4378
Qy	1759	GAAACAAGTACTGGTAATGACTCACAAAGTAGGCGCAGGAGATACAAAGTCGAGCCATAC	1818

D 4379 GAGCAGGTGAAAAATATAACATACGAGGAGCGCGGTTCACAGTGCAGCATAT 4438
Q 1819 CAGCGTAGGTGATTTGATACGAGAGCGCGGGTCCCTGTTCAAGCTTCAGGCATG 1878
D 4439 GACGGCAGGGTCTACTACCAATGTGATCGGCATTCGGTCCCTGAGTTTCAAGCTTGG 4498
Q 1879 AGTGAAGCGCTGATCGTTTTCACAGAGAGGAGTTCTGTAACAGATACCTGCACAC 1938
D 4499 AGCGAGAGCGGCATATGTGTGTACACGAAAGGAGTTCTGTAACAGAACTATACCT 4558
Q 1939 ATCCGATCAACGAGAGAGCGCTTAACACTGACGAGAGTACTATAGACTGTAAAGCT 1998
D 4559 ATGCGCTTCAAGCAGCGCTGCTGACACCCAGCAGAGGAACTACGAGAAAGTACAGCT 4618
Q 1999 CAGGACACAGCTGAGATATGCTGTGATATGACGACGAGAAAGTGTAAACGAGAA 2058
D 4619 GAAAGAACTGACGCGAGTACGTGTGACGTATGTAATAAAATCTGTGTCAAGAGAGAG 4678
Q 2059 GACGCAAGTCCCTTGTGCTTAACCGGTGATCTGTGATATCCACATTTCAAGATTGGC 2118
D 4679 GAACGTCGGGTGTTGGTGTGTGGAGAGCTAACCAACCCCGTTCAGATTTGGCC 4738
Q 2119 TACGAGAGTCTAAGACACGACGACGACCTCAACAAAGTCCAAACCATGAGTCTAT 2178
D 4739 TACCAAGGGCTGAAGATCAAGCGCTGCGCACATATAGACTACAGTAGTAGAGTCTT 4798
Q 2179 GGAATGCGAGTTGAGTAAATCTGAAATCATCAAAAGCGCTGATCTTAAGAAAGATCTG 2238
D 4799 GGGGTTCGGGATCAGGCAAGTCTGTATATTAAGAGCTGTGACCAAAACGATGTG 4858
Q 2239 GTTGTGATGCGAAGAAAGAACTGCGCAGAAATCATCAGAGATGTAAAGAGATGAGA 2298
D 4859 GTCAACAGCGGCAAGAGAGAACTGCCAGAAATATGTAACGATGTAAGAAAGACACGC 4918
Q 2299 CGATGATGTTGCTGCTAGAGTGTGATTCAGTCTTCTTAATGGGTTAAAGACCCC 2358
D 4919 GGGAGGGGACAAAGTAGGAAACAGTGAATCAATCTGTCTAAACGGGTGTGCTGTGCC 4978
Q 2359 GTTAAACCTGTATGATGATGAGGATTTGCTGCCATGCGAGGAGCGCTGTGCACTG 2418
D 4979 GTGACATCTATATGTGAGCAGAGCTTTCCTGTGCCATTCGGATCTCTGTGACCTTA 5038
Q 2419 ATTGCCATCTCAAACT---AAGAAAGTGTATTTGTCGGGGAGCCCAAAACAATGCGGC 2475
D 5039 ATTCTCTGTTAAACCTCGAGCAAGTGTGTATGCGAGAACCCCAACAATGCGGA 5098
Q 2476 TTCTTTAATGATGTGCTGTAAGATATTTTAACTGACATATGATGATGATGATG 2535
D 5099 TTCTTCAATATGATGAGCTTAAGTGAATTTCAACCAACATCTGCACTGAAGTATGT 5158
Q 2536 CATAAAGCATCTCTAGAGAGTGCACAGACTGTAACCGCATCTGTCCAGCTCTTC 2595
D 5159 CATAAAGTATATCAGACGTTGACCGCTCAGTCAAGGCGCATCTGTCTGACGTTGCAC 5218
Q 2596 TACGACAGCGAATGAAGACGTTAACCCATGTGCTGTAATAATCATATGATATGACA 2655
D 5219 TACGAGAGCAGATGTGCGACGACCAACCGTGCAAAACCATATATATGACACCA 5278
Q 2656 GGGACACAAAGCGCGACAAAGATGATGATTTCACTGTTTCAAGATGGGTGAA 2715
D 5279 GGAAGAGCAAGCCCAAGCAGAGACATCTGTGTTAATCATCTTCGAGGCTGGCAAG 5338
Q 2716 CAGCTACGATTTGACTAATAAATCAGAAATCATGACGCGGTGCACTGCCAGGACTT 2775
D 5339 CAGCTGAGTTGACTACGTTGACGAGATGATGACAGACGACATCTCAGGCGCTC 5398
Q 2776 ACGCGAAAGCGTTTATGCTGTGAGTAAAGTCAACGAGAAATCACTACTCGAG 2835
D 5399 ACCCGCAAGGGGTATGCGCGTAAAGGAGAAAGTGAATTCCTTTATGCCCT 5458
Q 2836 ACTTCTGACAGTGAACGTGTACTTACGACGACAGAAAAACGATTTCTGAGAGCG 2895
D 5459 GCGTGGAGCAGTGAATGTACTGTGACGCGCACTGAGATAGGCTGTGTGAAAAAG 5518

Q 2896 CTAGCTGTATCCCTGATTAAGACACTTACAGCTAAATATATCCGGGATTTCAAGCT 2955
D 5519 CTGGCGGCGATCTCCGTGATTTAAGTCTTATCAAACTTCCACAGGTAATTTACGCC 5578
Q 2956 TCATTTGACAGTGTGACGCGCAACAGACGCAATTAAGCAGCGCTTCTTATAGCCG 3015
D 5579 ACATTTGAAGATGTGCAGAAAGAACAGCAAAATATATGAAGTGTATTAAGACCGCT 5638
Q 3016 CAGACAGTATGTCTTCCAGATTAAGTAAAGTCTGTGGCGGAAGCTTTAGAACCA 3075
D 5639 GCGCTGTGACGCGTTCCAGAACCAAGCAAGCTGTGGGGAAGAACCTGTGCTCT 5698
Q 3076 GTCTTGCACGCGCAACATTTGTGCTGACAGACAGAGTGGGAGAC---GTTCACCCA 3132
D 5699 GTCTGACACTGCGGAATCAGATTAACAGAGAGAGTGGAGCACCATATATACAGCA 5758
Q 3133 TTCAAGCATGACAGAGGTAATCTGAAATGGAACCTGAACTTTTGTGACAGGTTTC 3192
D 5759 TTTAAGAGAGACAGAGCTTACTCTCAGTGTGCTTGAATGAATTTGACCAACCAATAC 5818
Q 3193 TTTGAGATGACCTGACAGTGTGTTATTTCCGCTCTACCGTCCGACTTACTTACAG 3252
D 5819 TATGAGTTGACCTGACAGTGTGCTGTGTTTCTGCCGAGGTGCTCTGTATTAAGAG 5878
Q 3253 GATCAGCACTGGGATTAATCTGCGCAGGAGAAACATGTATGGGCTTAATAGAGATGACA 3312
D 5879 AACCACTCTGGATTAACACACTGTGTGAAGATGTATGATTTCAATGCCGCAACGCT 5938
Q 3313 AAGAGTTGTACCGGCAATATCCGTGATCACAAAAGCGTTGACACAGCAGGATGACT 3372
D 5939 GCCAGCTGGAAGCTAGACATACCTCTGAAAGGCGAGTGTGCTAAGCGGCAAGCAGCA 5998
Q 3373 GATTTAAGAAATATATACATCAAGAGTACTCTTCAACATTAATGTGTCTTCAATTAAT 3432
D 5999 GTTATGCAAGAAAGAAATCAACCGCTTCTGTGTGACAAATGTAATTTCTATCAAC 6058
Q 3433 CGCCGTTGCCCACTGTTGATGTGTGACCAACAAAGACAGGTAACATGATCAGC 3492
D 6059 CGCAGCTGCCGACCGCTGTGTGTGATGACAAAGACGTTAAAGCAATGAGGTTGAG 6118
Q 3493 GGATTCCTATTAAGATGAAGGCAATCTGTGTGTGATGCGGATCTTATCAGATT 3552
D 6119 TGCTGTGATTAAGATGAAGGTTACCACTCTCTGTGTGATGATCAACCTGCT 6178
Q 3553 CCAGGAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3603
D 6179 TTGCTTGAAGCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6238
Q 3604 TGTATCTGATTTGGGAATACCTAGGCAATGTCGGTAAATATGATCATATCTTTGCAAT 3663
D 6239 TACGACTTAAGTTAGACGTCGCGGTGACGCGGAGGTTGACATTTGTTGTGAC 6298
Q 3664 GTTGAAGCCCGTACAGAAACATCACTACCAAGTGTGAGATGACGCTATCACAC 3723
D 6299 ATTCAACGGAATTCAGAAATCAACATCAACAGAGTGTGTCGACACGCGCATATAC 6358
Q 3724 AGCATGTAAAGTGTAAAGTGTGTCACCACTGAACATGCGGGAACATGTGTGCTATA 3783
D 6359 CAGATGTGTGGGAGATGTGCTACGACTGTAAACCCGCGGCACT---TGATGAGA 6415
Q 3784 GGGTATGGGCTTGTGATTCGCGCAACGAGAAATTCATCACTGCGGTGCGACGCTCAAT 3843
D 6416 GCTTACGATATGCGCATTAATATAGGAAAGCGTTTCTCTTCAAGAGAAATTC 6475
Q 3844 AGGTTTACCGGTGTCTGTACGCTTACAGAACCTGCGGAAATATCTAGGTTCTTCTG 3903
D 6476 TCGTCTCAAGAGTGTGCGCGGATTTGTGTCAACGAAATATCAGAAAGTGTCTG 6535
Q 3904 TTCTTGGCAAGACAAACGCGCAACACATATGACAGAGACATCTCGGTGTGATGCTT 3963
D 6536 TTCTCCAACTTTTGAACGAGAAAGAACCTCTTACGCTTACACGATGAATTAACCAAGCTG 6595

3964 GACAACTATCAAGGTCACCAAGGTCAGAGGAGGAGGAGGTCACAGAGT 4023
Db AGTCCGCTATGCGCGAAGACCAATGACACGCGCGGCTGTCACATCTCAAGAGT 6655
Qy ATCAGAGGATATAGCAAGGCGCTGACCAAGCTATCTGTAATGCTCTAATAGCAA 4083
Db AAGAGAGCAGATAGCCACGTCACAGAGCGGCTGTGTTACGACAGTAAACCGCT 6715
Qy GGTCAACCAAGTCCGAGTGTCCGCTGACATGTAACCAAAATGCGCGCTCTTTGAT 4143
Db GGAATGTAGAGGAGATGCGCTATGCAAGGCGCGTGGGGAAGAAATGCGCGCTTTAG 6775
Qy AGACACCAATATGCTGTGGGACGCGCTAGACTTTGTAAGCAAGAACCG---CTCATATA 4200
Db GGAGCAGCAACACAGTGGGACAAATTAACAGTATGTGGCGCTCGTACCCGCTATC 6835
Qy CATGCTATAGCAACCAATTTTCTAAGATCCGGAACCGGAGGAGGAGCTTAAGCTCGCA 4260
Db CACGCTGTAGCGCTAATTTCTCTGCAAGCTGAAGCGGAAGGAGCGCAATTTGGCC 6895
Qy GCTGCTCATATGAGCATAGCGCTCATCTGCAAGCTGAGCGGATTAACAAAATATCAGTA 4320
Db GCTGCTACCGGCGAGTGGCGCCCAAGTAACAGACTGTCACTGACGAGCTACCATC 6955
Qy CCGCTACTGTCAACCGGCACTATTTCTGTGGCAAGATCGAGTATGCAATCATTTGAT 4380
Db CCGCTGTCTCCACAGAGTGTTCAGCGCGGAAGATAGCTGACAGAAATCCCTCAAC 7015
Qy CACCTGTCTACTGCTTTGACACTAGCGATGCCGATGTCACATATATTTGCTTGATAA 4440
Db CATCTATTCACAGCAATGACCGCAAGACGCTGACGATGATCTACTGACAGACAA 7075
Qy CAATGGAGACCAAGATATCGAGGCAATTCACCCCAAGAAAGCGTAATTTGAT 4500
Db AGTTGGAGAGAAATCCAGAAACCATTCAGAGAGCGCTGTGAGTGTCTCAAT 7135
Qy GATGCAAGCCAGTATGACATTTGCTGAGGTCACCCCAAGACGCTTTTGGCAGGC 4560
Db GATGACGTGAGACTACCAAGACTTTGGTAGAGTGCACCGGACAGCGCTGTGGGT 7195
Qy AGACCAAGTACTCGCTCAATGAGGCAAGTTGTTTCACTACTGGAAGTACAGATTC 4620
Db CGTAGGCTACAGTACCACTGACCGGTCGCTGTACTCTTTGAAAGTACGAAATTC 7255
Qy CATGACCGGCAAGGATTCGCGAAATTCATGCAATGTGGCCCAACAAATCTGAGGT 4680
Db AACCGGCTGCTATGATGACAGATGCTGACGTTGTGGCCAGACTGCAAGAGGA 7315
Qy AATGACGATTTGCTGTATCATCTGTGGGAGAGTATGTCAGCATCCGCTCCAAATGC 4740
Db AACGAAAGATATGCTATAGCGCGCTGGGCGAAACAAATGCAATGATCCAAATGT 7375
Qy CCAATGAGAGTCAAGAGCGCTGTCTCACTCAACAACCTTCATGCTGTGTAATTAC 4800
Db CCGGTAACGATTCGATTCATCAACACTCCCAAGAGAGTCCCTGCTGTGCGCTAC 7435
Qy GCTATGAGGCTGAGCGGTATACAGTGTGCGCTCTGGAAGAAAGACAGTTCGCTTA 4860
Db GCAATGAGAGCAAGGAGATGCGCGCTCTAGGTACACCAAGTTAAAGCATGTGTGT 7495
Qy TGCTCATCATCTGTTGCGGAGTACAGATCAAGCGCTGCAAGACTACAGTACAGC 4920
Db TGCTCATCTTTTCCCTCCGAAATACATGTAATGAGGTGAGAAAGTACAG 7555
Qy AAACCAAGTCTGTTTCAAGCGCTGCTACCAACGCTGTACACCCAGAGAAAGTACGCGAA 4980
Db AAGTTCTCTGTTGACCCGAGCGTACTTCAAGGTGTAGTCCGCGAAGATATCCGCA 7615

GENERAL INFORMATION:
APPLICANT: Parrington, Mark
APPLICANT: Li, Xiaomao
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
FILE REFERENCE: 1038-1042 MIS
CURRENT FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/065,791
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: PCT/CA98/01064
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 8100
TYPE: DNA
ORGANISM: respiratory syncytial virus
US-09-554-337-4

Query Match 29.5%; Score 1473.8; DB 4; Length 8100;
Beet Local Similarity 57.9%; Pred. No. 0; Mismatches 1982; Indels 36; Gaps 8;
Matches 2779; Conservative 0; Mismatches 1982; Indels 36; Gaps 8;
Qy 214 ATTGAAAGTGGCGCCGTCGACATGACATTCACATTCACCGCTATCATTTGATGCGCT 273
Db 1 ATCGCATGTGGCTCTTCAGAGAGATGATGTCTACGCAAAATACCATGCTGCTATGCGCT 60
Qy 214 ATGTAAGCGCTGAAGACCGGACAGACTACACGATACAGGATATGCAAGAAAGCTTAAGAAA-- 331
Db 61 ATGCGCAGCGGAGAAAGCCCGAAAGGCTCATATGCTACGAAAGAACTGGACAGCGCC 120
Qy 332 ----GTGACATTACCGAAGAAACATAGCCCTAAGCGGAGAGCTGTGGAAGTCTT 387
Db 121 TCCGGAAGGTGCTGATGAGAGATCCAGAGAAATCACCGACTGACAGCGCTAC 180
Qy 388 TCAACACCGAGCAGAGACTCCATCTGTGTATGACACAGACGCGCAGTGTAGTAC 447
Db 181 GCTACCGCAGAGCGTGAATCTCTACCTTTGCTGCAATACAGAGCTACCGTGTAG 240
Qy 448 TTGGAAAGTATGACATATCCAAAGATGTGTACGAGTCCATGCAACGACATCAATCTAC 507
Db 241 GCAGCCGAAGTGGCGCTATACAGAGACGTGTATGCTGTATGACACCAACATCGCTGTAC 300
Qy 508 CACACCGCGCTTAAAGAGTATGACACATTTACTGAGATAGCTTTGACAGACCCCTTT 567
Db 301 CATAGCGCATGAAGAGTGTGAGAACGCGGTATGTGATGTGGTTGACACACCCGCTT 360
Qy 568 ATGTACAAATCATGCGAGTTCCTTCTACTTACCAACAGAACTGGGCTGACAGAGA 627
Db 361 ATGTTTACCGGCTAGCAGCGCGCTATCTCAACCTTACGCAAACTGGGCGGACAGCAG 420
Qy 628 GTATTGAAAGCAGTAACATTTGCGCTGTACTAGATCTTCAAGAGAGCAGGCTTGA 687
Db 421 GTGTACAGGCGCAGAGACATATGACTGTGTGACAGATCTTGAATGAGGAAAGTAC 480
Qy 688 AAATCTCAATCTTGAAGAAAGAGGCTCAACCTACTAATTAAGATATATTTCTCGGT 747
Db 481 AAATGCTCATCTTCCGGAAGAACATTTGAACCTTGCACACAGATATGTTCGGTA 540
Qy 748 GGTTCACAAATCTACAGAGATAGATCACTGTATAGTGTGATCTTCAAAAGG 807
Db 541 GGAATCTATATGTACCTGAGAGCAAGAGTACTGAGAGACTGCACTTACCTTCGTA 600
Qy 808 TTCACTTGAAGAAAGTCTAATCTCAACAGTATGTGAGACATTTGACGCTGTGA 867
Db 601 TTCACTTGAAGAAAGTCAATCTTATCTGTAGTGTGAGTACATATGTGTA 660
Qy 868 GGTATGCTATCAAAAGATATAGATACGCCAGACATATCGTAAAGTTGAAAGCTTG 927
Db 661 GGTATGATTAAGAAATCTATGTGCTCCGCGCTGTAGCGTAAACGTTAGGCTAC 720
Qy 928 GGTTCACAAATGATCGGAGGCTTCTTGTAGTTGCAAAAGTACAGATACGCTGCGCGC 987

Db	721	GCCTGACGTAATCAACGCGGAAGGATTTCTTAGTGTGCAGAACCAACGACATCTGTCAAAAG	780
Qy	988	GAGAGGGTTTCTTTTGCTGTGTGTACGTATGTACAGGCACACTTTGGCATCAGATGACA	1047
Db	781	GAAGAAGCTCATTTCCCTGTATGACCTACAGTCCCTCCATCAACATCTGTGATCAATGACT	840
Qy	1048	GGGATTTCTGGCACTGACGTTAAGTGTGATGACGCAAAAACATAATTGGTGGCTCAAC	1107
Db	841	GGCATACTAGAGCACCGACGTACACCCGAGAGACGACAAAGTTGTAAGTGGATTGAT	900
Qy	1108	CAAAGGATGTGCTCAATGTATAGGACGCAAAAGAAATATACTACAAATGCAGACATATCTA	1167
Db	901	CAGAGGATAGTTGTGAAACGAAAGAACACAGGAAACATAACGATTAAGAACTATCTG	960
Qy	1168	TTACCAGTGTGTGCGCCACGACGCTTTTCCAGTGTGGGCGCTGGAACATCTGTCCGACTTGGAC	1227
Db	961	CTTCGATGTGTGGCCGTGTGGCATTTTAGCAAGTGTGGGAGAGAAATACAAAGCAGACTTGAT	1020
Qy	1228	GACCAAAAGAATAAGGGGTGCGGAGCGCACTTTACTATGGCTGTCTGTGGCTTTC	1287
Db	1021	GATATAAAACCTGTGGGTGTCCGAGAGGTGCATTACTTGCTGTCTGTGTGGGCAATTT	1080
Qy	1288	AAGACCCAGAAATATCAATCCATCTACAGAAAGCTGTAGCGAAACATTAAGAAATA	1347
Db	1081	AAACGAGGAAGATGCACACATGTACAGAAACAGACACCCAGACAAATGTGAAGTGTG	1140
Qy	1348	CTGCGCGTCTTTGACTCATTTTGTGATTTCCAGCCTTACCAACCAACGCGGCTCGATATGGGC	1407
Db	1141	CTTTCAAGGTTTAATCTGGTGTGTATCCGAGCCTATGTGTACAGGCTGTGCAATCCCA	1200
Qy	1408	TTCCGCGGTAGGCTCAGCTGCTGCTTGAACCAACTGTCAAAACCGCACCGGCTATTACA	1467
Db	1201	GTCAAGATCACCCATTAATAGATCTTTTGGCCAAAGAACCAAGCGAGA---GTATAATCTT	1257
Qy	1468	ATGGCCGATGTGAGACATCTGTGTGGCTTACAGCAAGAGCTGAAGATGGCTGCACGC	1527
Db	1258	GTTCTCAGCGGTGTGTGACGAGGAGTCTGAACAAGAGAGAAAGAGAGTTGGAGCC	1317
Qy	1528	GAAGAGATCAGAGAAAGCCCTCCACCCCTGTGCTCCCTGAATAAGAAAAAGAGACGATAGAG	1587
Db	1318	GAGCTGAATAGAGAAAGCTTACACACCCCTGTGCTCCCATCGGCGCGGACGAGACGGAGTCT	1377
Qy	1588	GCAGAGTAGACCTTCATTATATGCAAGAG-----GCAGGAGAGGTAGCTGTGAGACACCA	1641
Db	1378	GTCAAGCTGCAGCTTGAAGAACTAGATGATACGACGAGGTGACGGGCTGTGGAACACCTT	1437
Qy	1642	CGAGAGACATCAGGGGTGACAAAGTTTACCAAGCCAAAGAAAGATTGGTCTTAACGCTATA	1701
Db	1438	CGAGCGCGTTGAAGATCACCGCACAGCGCAACGACGATCTATAGGAAATTAACGTAGTTT	1497
Qy	1702	CTTTACCCCGAGCGGTATTAATATGTAAGAAAACTGTGCTGTATTCACCCATTGGCGGAA	1761
Db	1498	CTGTCCCGCAGACCGTGTCTAAGAGCTTCAAAGTTGGGCCCCCGTGCACCTCTAGACAGAG	1557
Qy	1762	CAAGTACTGTATGACTCACAAAAGTAGAGGCAAGGAGATACAAAGTGTGACCATACAC	1821
Db	1558	CAGGTGAATAATATACATATACGAGGAGGCGCGGTTCACAGGTGCAACGATATATAC	1617
Qy	1822	GCTTAAGTTCATTTGATCCGAAGAGGACCGCGGTCTCTGTCAAAGCTTCCAGCATTTGAT	1881
Db	1618	GGCAGGGTCTCACTACATCAATGTGATCGGCATTTCCGAGTCCCTGAGTTTCAAGCTTTGAGC	1677
Qy	1882	GAGAGCGCTAGATCGTTTTCAACGAGAGGGAGTTGTAACAGATACCTGCACACACATC	1941
Db	1678	GAGAGCGGCATTAATGTGTATACAGAAAGGAGTTTGTCAACAGGAACTATACATATTT	1737
Qy	1942	GCAATCAACGAGAGACGCTAAACACTGACGAAGATCTATAAGACTGTAAAGACTCAG	2001
Db	1738	GCCGTTCAACGAGACCGTGTGCTGAACACGACGAGGAGATCTACGAAAGTCAAGACTGA	1797
Qy	2002	GACACAGACTCAGATATGCTTGTGATTTTGACGCAAGAAAGTGTATTACCGAGAAAGC	2061

D	b	1798	AGAACTGAACCGCGAGTACGTTTCGACGTAGTAAATAAATGCTGCGTCAAGAGAGAGAA	1897
O	y	2062	GCAGGTCCTTGTGTCCTPAACCGGTGATCTGGTAGTCCACCATTTCACGATTTGCGTAC	2121
D	b	1858	GCCTCGGGTTTGTGTGTGGTGGAGACCTPAACCAACCCCGCTTCCATGAAATTCGCTTAC	1917
O	y	2122	GAGAGTCTCAAGACGACGACGACGACCTCACAAATCCCAACCATCGGAGGTATGGA	2181
D	b	1918	GAAGGGCTGAAGATCAAGCGCGTGGCACCATATPAAGCTACAGTAGTAGGAGTCTTGGG	1977
O	y	2242	GTAGTGCAGAAAGAAAACTGCGCAGAAATCATCGAGAGTGTAAAGAGATGAGAGT	2301
D	b	2038	ACGACGGCAAGAAAGAAAGAACTGCCAGAAATAGTTAAACGACGTGAAGAAAGCACCGCGGG	2097
O	y	2302	ATGATATGTGTGCTAGAGACTGTGATTCAGTGTCTTAAATGGGGTTAAGACCCCGTT	2361
D	b	2098	AAGGGGACAAAGTAGGGAAAAACAGTACCTCCATCTGCTTAAACGGGTGTCGTGCGGTG	2157
O	y	2362	AACACTCTGTATCATGTATGAGGCACTTTGGCCATGCTACAGGAGCGCTGTGACACTATT	2421
D	b	2158	GACATCTCTATATGTGGAGCGAGGCTTTGGCTTGCATTCGGTACTCTGTGCGCTTAAT	2217
O	y	2422	GCCATCTGTCAAACCT--AAGAAAGTGTATTTGTGCGGGGACCCAAACAAATGCGGCTTC	2478
D	b	2218	GCTCTGTGTAAACCTCGGAGCAAAAGTGTGTATGCGGAGACCCCAAGCAATGCGGATTC	2277
O	y	2479	TTTAACTGATGTGCTCGGAAGTACATTTTAAACATGACATATGCACTGAAGTGTACAT	2538
D	b	2278	TTCAATATGATGACGCTTAAAGGTGAACCTTCAACCAACATCTGCACTGAAGTATGTAT	2337
O	y	2539	AAAGCACTCTTAGAGAGTGCACACAGACTGTAAACCCCATCTGCTCACGCTCTTCTAC	2598
D	b	2338	AAAAGTATATTCACGACGTTTGCACGCGTCCAGTACCGGCATCGTGTCTACGTTGACATAC	2397
O	y	2599	GACAAAGCAATGAAAGACGCGTTAACCCATGTGTGATTAATCATCATATGATATCCACAGG	2658
D	b	2398	GGAGGCAAGATGCGCCACGACCAACCCGCTGCACAAACCATTAATCATATGACACACAGGA	2457
O	y	2659	ACCACAAAGCCGCAACAAAGATGATCTGTATTTCTTCAAGAGATGGGTGAACAG	2718
D	b	2458	CAGACCAAGCCCAAGCCAGAGACATGTGTAAATCATGCTTCCAGGCTGGGCAAAACAG	2517
O	y	2719	CTACAGATTGACTCAAAAATCAAGAAATATGATCTGCGGTGATGCGCAAGACTTACG	2778
D	b	2518	CTGCAATTGACTACCGTGACACGAAAGTCAATGACAGACAGACATCTCAGGCGCTCAC	2577
O	y	2779	CGAAAGGCGCTTATGTGCTCAAGGTACAAAGTCAACAGAAATCCACTACTCGACACT	2838
D	b	2578	CGAAAGGGGTATACGCGCTPAAGCCAAAGGTGAATGAATAATCCCTGTATGCCCTGCG	2637
O	y	2839	TCTGAGCACTGAAACGTGTATCTTACACGCAAGAAAAAGCATTTGCTGGAAGCGCTA	2898
D	b	2638	TGCGAGCAGTGAATGTACTGCTGACCCGCACTGAGGATAGCGTGTGGTGAACACGCTG	2697
O	y	2899	GCTGGTATCCCTGAGATTAAGACACTTACAGCTAAATATCCCGGGGATTCACGGCTTCA	2958
D	b	2698	GCCGGCGATCCCTTGATTTAAGGTCTTATCAACAATTCACAGGTTAACTTTAGCGGCACA	2757
O	y	2959	TTGAGCAGACTGCGACCGCAACAGACGCCATTATGCGACGCGTTCTTTGATTAAGCCGCA	3018
D	b	2758	TTGGAAAGATGGCAAGAAAGAACAGACAAATATATGAAGGTATGAAGACCGGCTGCG	2817
O	y	3019	ACAGCTGATGTCTTCCGAATPAAGGTAAAGTCTGCTGGGCGAAGGCTTTAAGACCACTC	3078
D	b	2818	CTGTGACGCGTGTCCGAACAAAGCCAAAGCGTGTGTGGGCGAAGAACCTGTGTGCTCTC	2877
O	y	3079	TTGGCCACCGCAACATTTGTGCTGACAGACAGAGTGGGAGAC---GTGGACCCATTC	3135
D	b	2878	CTGGAACCTGCGGAACTCAATTGACGCAAGAGAGTGAAGACATATTAACGACATTT	2937

QY 3136 AAGCATGACAGACCGTACTCACTGAAATGSCACTGAACCTCTTTTGACACAGTTCTTT 3195
 DB 2938 AAGGAGACAGAGCTTACTCTCAAGTGTGCGCTTGAATGAAATTTGACCAAGTACTAT 2997
 QY 3136 GAGATGACCTGACAGTGGGTTATTTCCGCTCTCAACCGTCACTTATTAACAGGAT 3255
 DB 2998 GAGATGACCTGACAGTGGGTTATTTCCGCTCTCAACCGTCACTTATTAACAGGAT 3057
 QY 3256 CAGCATGGGATTAACCTCGCAGGAAAGACATGTAATGGCTTAATAGAAAGTAGCAAG 3315
 DB 3058 AACCATGGGATTAACAGACTGATGAAAGATGTAATGATTCACCAACAGTGGC 3117
 QY 3316 GAGTTGTCAGGCGATATCCGTGATCACAAAAACGGTTGACACAGGAGGATGAT 3375
 DB 3118 AGGCTGGAAGTGAACATCTTCTGAAAGGGGCACTGCAATCGGCAAGACGATTT 3177
 QY 3376 ATAGGAATATATCAATCAAGGACTATCTCCAAATTAATGTGTTCCATTAATCCG 3435
 DB 3178 ATCGAGAAAGAAATCCAAACCGCTTCTGTGTGACAAATGTAATCTTATCAACGCG 3237
 QY 3436 CGGTTGCCCACTGCTGATCGTTGACCAAAAGACAGGATCAACTATCAACGGA 3495
 DB 3238 AGGCTGCGCAGCGCTGTGTGCTGATCAAGACGTTAAAGCGATGAGGTTGATG 3297
 QY 3496 TTCTATCTAAGATGAAGGCAAAATCTGTGTGATGCGGATCTTATCAGATTCGA 3555
 DB 3298 CTGTGATTAAGTGAAGGATCAAGCTCTGTGTGATGATGATGATCAACCTGGCTTG 3357
 QY 3556 GGGAAAGATGAGTCCATGAGTCCATTCCTCACTATAC-----CATCAGTGT 3606
 DB 3358 CTTGACGAGGATCACTGTGTGTGTCACCGCTGAATGTCAAGGCGCGATAGTGTAC 3417
 QY 3607 GATCTCATTTTGGGAATACCTAGCCATGTCGTTAAATGACATTAATCTTTGTCAATGT 3666
 DB 3418 GACCTAAGTTTGAAGTGTGCGCTGACGCGGCAAGTTGACTGTCTTTGTGAACAT 3477
 QY 3657 AGGACCCGTCAGGAACCATCACTACCAAGTGCAGATCAAGCTATCCACACAGC 3726
 DB 3478 CACACAGGAATTAAGATTCACACATACAGCAAGTGTGTGACACGCAATGAGCTGACG 3537
 QY 3727 ATGCTAAGCTGTAAGGCTGTCCACCACTGAACACTGCGGAAACATGTGTGCTATAGG 3786
 DB 3538 ATGCTTGGGGAGATGCGCTGACAGTCTAATAACCGGCGCATCT---TGATGAGAGT 3594
 QY 3787 TATGGCTTGTGATGCGGCAACCGAATATCATCACTGCGTGACGCTCATTTAGG 3846
 DB 3595 TACGATACGCGATTAATCAAGCAAGCGTGTCTCTCTTAAGCAAGATTTCTG 3654
 QY 3847 TTACCCGCTGTGTCAGCTTAAGAACATGCGGAAATATGAGGTTCTTCCGTTTC 3906
 DB 3655 TCTGAAAGTGTGCGCCGATGTGTGTCACACGAAATACAGAGTGTCTTGTGTTTC 3714
 QY 3907 TTCCGCAAGACCAACGCAACCAACATCAACATGACAGACGATGATGCTTGAC 3966
 DB 3715 TCCAACTTTGAACGAAAGAGACCTCTAGCTAACACATGAAATCAAGCTGAGT 3774
 QY 3967 AACATCTATCAAGGCTCAACAGGATCAAGGCAAGGAGCTTCAAGCTATCAAGTATC 4026
 DB 3775 GCGGTGTGCGGAAAGCCATGACACGCGGCTGTGCAACATCTCAAGAGTTAG 3834
 QY 4027 AGAGTGAATTAAGAAAGGCGCTGACCAAGCTATCTGTAATGCTGTAATAGCAAGT 4086
 DB 3835 AAGACAGACATACCACTGTCACAGAAAGCGCTGTGTGTAAGCACTTAACCGCGTGA 3894
 QY 4087 CAACAGGTTCCGAGTGTGCGGTGATGATCCGAAATGCGCGGCTTTGTATGA 4146
 DB 3895 ACTGTAAGGGAATGGGTATGCAAGGCGGTGGGAAAGAAATGCGCTTAAAGGA 3954
 QY 4147 CAGCCATATGCTGTGGGAGCGCTGACCTTGTGAAGACGAACG---CTATCATCAT 4203
 DB 3955 GCAGCAACCAAGTGGGCAATTAACAGTCAATGTGCGCTCATACCCCTCATTCAC 4014

QY 4204 GCTGTAGGACCCCAATTTTCTAAGATGCGGAAACCGGACGCTTAAGCTCCAGCT 4263
 DB 4015 GCTGTAGGACCCCAATTTTCTTCCACGATGGAAGGGAACCGGAATTTGCGCGT 4074
 QY 4264 GCTTACATGACATAGCGTCCATCGTCAACGCTGAGCGGATTAACAAAATATCACTACG 4323
 DB 4075 GTCTACCGGACAGTGGCGCGCAAGTAAACAGACTGTCACTAGAGAGCTACCTCCCG 4134
 QY 4324 CTACTGTCAACCGGATCTATTTCTGTGCGCAAAAGTTCAGATATCAATCATTCATC 4383
 DB 4135 CTGTGTCCACAGAGTGTTCAGCGCGGAAAGATAGGCTGCACACATCTCTCAACCAT 4194
 QY 4384 CTGTCTACTGCTTTGACACTTACGATGCGGATGTCACCATATATTTGCTGATTAACA 4443
 DB 4195 CTATTCAAGCAATGAGAGCCGACGGAACGTGACATGACATCTTACTGCAAGACAAAAGT 4254
 QY 4444 TGGAGACAGATTAATGAGGCAATTCACGCAAAAGAGCGTCAAAATTTCTGATGAT 4503
 DB 4255 TGGAGAAAGAAATCCAGGAAGCCATTCAGATGAGGACGCGTGTGAGTGTCTCAATGAT 4314
 QY 4504 GACAAAGCAATGATGATCTTGTGAGGCTCCACCAACAGCTCTTTGACAGCAGA 4563
 DB 4315 GACGTGAGCTGACCAAGACTTGTGTGAGAGTGACCCGACAGCAGCCTGTGTGCTGT 4374
 QY 4564 CCAGGTTACTCCGTCATGAGGCGCAAGTTGATTAATCACTGGAAGTACACGATTCAT 4623
 DB 4375 AAGGCTTACAGTACCACTGACGCGTGTGCTGTACTGTGTAAGGATGCAAAATTCAC 4434
 QY 4624 CAGACCGCCAAAGACATTTGCGGAAATCCATGCAATGTGCGCCCAAAATCTGAGCTAAT 4683
 DB 4435 CAGCTGCTATGATGATGAGAGATGATGAGATGATGAGTGTGCGCCAGCTGCAAGCAAAAC 4494
 QY 4684 GAGCAGATTTGCTTGTATCATCTGTGGGAGAGATATGTCAGATTCGCTCCAAATGCGCA 4743
 DB 4495 GAAACGATATGCTATGACCGCTGTGGCGAAACAAATGGAACATGATCCAAATGTCCG 4554
 QY 4744 GTAGAGATGACAGGCGCTGCTCACCTGACACACTTCCATGCTGTGTAATAGCT 4803
 DB 4555 GTGAACGATTTCCGATTCATCAACACTTCCAGAGACAGTGCCTGTGCTGCGCTACGCA 4614
 QY 4804 ATGACGGGTGAGCGGATTAACAGGTTGCGCTTGTGGAAGAAAGAACAGTTGCCGTATGC 4863
 DB 4615 ATGACAGCAGAAACGATGCGCGCTTGTAGTTCACACCAAGTTAAAGATGATGCTTGC 4674
 QY 4864 TCATCATCTCTGTTCCGGAATGACAGATCAACAGCGGTGCAAGCTACAGTGCAGCAAA 4923
 DB 4675 TCATCTTTTCCCTCCCAAAATACCATGTAGATGAGGAGTGCAGAAAGTGAAGTGCAGAAAG 4734
 QY 4924 CCAGTCTGTTTTCAGGCGGTGATCACCGGCTGTACACCCAGGAAGTACGCGGA 4980
 DB 4735 GTTCTCTGTTTCAGCCGACGATCTTCAATGATGATGATGCGGAGATGATGCGCA 4791

RESULT 8
 US-08-446-932-1
 ; Sequence 1, Application US/08446932
 ; Patent No. 5639650
 ; GENERAL INFORMATION:
 ; APPLICANT: Johnston, Robert E.
 ; APPLICANT: Simpson, Dennis
 ; APPLICANT: Davis, Nancy J.
 ; TITLE OF INVENTION: cDNA Clone for South African
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kenneth D. Sibley
 ; STREET: Post Office Drawer 34009
 ; CITY: Charlotte
 ; STATE: NC
 ; COUNTRY: USA
 ; ZIP: 28234
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,932
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-118
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 881-3140
TELEFAX: (919) 881-3175
TELEX: 575102
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-446-932-1

Query Match 26.7%; Score 1333.2; DB 1; Length 11663;

Best Local Similarity 56.1%; Pred. No. 0; Mismatches 2128; Indels 80; Gaps 13;

Matches 2820; Conservative 0;

23 ATATGAAAGAAATTCAGCTTACTTGAAGTCTGACAGCCCTATGTCAGAGTCTTAAAGC 82
Db AGAAGCAGATGTTACGTAGACGTAGACCTCAGAGTCCCTTGTGTGTAAGTGAAG 123
QY 83 GAGAGTTTCCAAATTTGAGATCGAAGCAGAGTCTGACATGACATGACCTCCATG 142
Db AGAGCTTCCGCAATTTGAGATGACAGACAGAGTCTGACATGACATGACCTCTATG 183
QY 143 CCAGAGGTTTGGCAGTGGCAACAACTTGAAGAGAGTGAAGCCGGAGCCAG 202
Db 184 CCAGAGATTTCCGATCTGCGAGTAACTAATGAGCTGGAAGTCTCTCAACAGGA 243
QY 203 TTATCTTGAACATTTGAAGTGGCCGCTGACAGATGACATTTCCATTCACCGTATG 262
Db 244 CGATTTTGAACATGAGGAGCGGACCGGCTGTAGATGTTTTCGAGAGCAGATGACAT 303
QY 263 GTATCTGCTTATGATTAAGCCTGAAGCCGGAACAGATCAACAGTATGCAAGAAC 322
Db 304 GCGTTTCCGCTATGATGTCAGAAAGCCGGAACCGATGATGAATATGCGCAAAAC 363
QY 323 TTAAAGAAA-----GTGACATTAACGCAAGAACTATAGCCTTAAGCCGCGACCTGC 376
Db 364 TGGCGGAAAAGCATGTAAGTAAACAAAGAACTTGCATGAGAAAGTCAAGAACCTCC 423
QY 377 TGAAGTATGATCAACACAGAGCCAGAGACTCCATCTGTGTATGACACAGACCA 436
Db 424 GAGACGATCTTGAATACACCGGATGCTGAAGCCATCATCTGTTCCAGACAGATGTA 483
QY 437 CGTGTAGTACTTTGAAAGTGTAGAGATTAACAGATGTGTACGAGTCCATCACCGA 496
Db 484 CTGTGCAACAGCGGTGCGAGTACTCGTCATGACAGAGCTGTACAAA--TCAACCTCCG 540
QY 497 CATAATCTTACACAGGCGCTTAAAGAGTTAAGCAATTTAAGTATAGCTTTTACA 556
Db 541 GAATATTTTACACAGGCTTAAAGAGCGTGCAGCCGTGTATGATGAGTTGCTTCA 600
QY 557 CGAGCCCTTTTATTAACAAAACATGCGAGGTTCTTACCTTACTTAACACAGAACTGG 616
Db 601 CCACCCAGTCTATTTCTCGGCTATGCGAGTTGTGACCTGCAACACCAACGAGTGG 660
QY 617 CTGACGAGAGATTTGAGAGCAGTAACTTGGCTTGTGTAATCTCAAGATTTTCAAGAGA 676
Db 661 CCGAGCAAAAGTCTCTGAAGCGCGTAACTGAGACTCTGACGACAAAAGCTGAGTGAAG 720

QY 677 GCAGGCTTGGAAAACTCTCAATCTTGAAGAAGAGGCTCAACCTACTAATATAGATCA 736
Db 721 GCAGGACAGAAAAGTTTCTCATATATGAGAAAGAGTTGAAGCCCGGCTACGGTTT 780
QY 737 TATTTCTGGTTGGTTCAACAATCTACAGAGATATGATCATGTTAGCTGGCATC 796
Db 781 ATTCTCTGGATGACACTTTTACCAAGAACAGAGCCAGCTTGCAGAGCTGCGATC 840
QY 797 TTCCAAACGTGTTCCACTTGAAGAGAAAGTCACTTCAAGATGATGAGGACATTTG 856
Db 841 TTCCATCGTGTTCATTTAAAGGAAAGAGTGTACCTTCCGCTGTGATCACTGG 900
QY 857 TCAGCTGTGAAGGATAGCTATCAAAAAGATTAAGATCAAGCCAGAGCTATACGTTAAG 916
Db 901 TGAGCTGCAAGGCTAGTATGTAAGAAATCAACATCATCTCCCGGATACCGGAGAAA 960
QY 917 TTGAGAACTTGGCGTCCAAATGATCGAGAGGTTTCTTGAAGTTGCAAGTCAAGATA 976
Db 961 CCGTGGGATACGGGTTTAAACAAATAGCAGAGGCTTCTTGTATGCAAGTTAACGATA 1020
QY 977 CGCTGCGCGGAGAGGTTTCTTTTGTGTGTATGATGTATGACCAAGCCACTTGG 1036
Db 1021 CAGTAAAGAGAAACGGGATATGTTCCCGTGTGACAGATATTCGCGCCACATATGCG 1080
QY 1037 ATCAGATGACAGGATTTCTGCAACTGACCTTATGATGATGACGACAAAACCTATTGG 1096
Db 1081 ATCAGATGACCGGATATATGACAGGATATCTCAGTGAAGATGACAAAACCTTTGG 1140
QY 1097 TTGGGCTCAACCAAGAGATTTGCTCAATGATGAGAGCAGAAAGATTAACAAATGC 1156
Db 1141 TTGGGCTCAACAGCGAATGCTCATTAACGTTAAGACTAACAGAAACAAATATCATGC 1200
QY 1157 AGAATCTATTAATCAAGTGTGCGCCAGGCTTTTCCAGTGGGCGCGTGAACATGCTG 1216
Db 1201 AAAATTAATCTTCCCAATATATGCAACAGGTTCAAGAAATGGCCCAAGAGCGCAAG 1260
QY 1217 CCGACTTGAAGCAGAGAAAGAACTAGAGGTTGCGGAGCCGACTTACTATGAGGCTGCT 1276
Db 1261 AAGATCTTGAATGAAAAATGCTGGGACAGAGAGCCGAACTTATGATGCTGCT 1320
QY 1277 GCTGGGCTTCAAGACCCAGAAAAATCAATCATTAAGAAAGCTGTATGCAACAA 1336
Db 1321 TGTGGGCGTTTCCGACTAAGAAAGTGCATCTTATTCGCCACCTGGAACAGACCA 1380
QY 1337 TTAAAGAAAGTACTCGGCTTTTGAATCTTATGATGATTCACGCCCTTACAGCCAGCG 1396
Db 1381 TCGTAAAGTCCAGGCTCTTTTAAAGCGTTTCCCATGTCATCGTATGACTTACTT 1440
QY 1397 TCGATATGAGGCTTCCGCGTAGGCTCAAGCTGCTGTGAACCA---ACTGTCAACCCG 1453
Db 1441 TGCCCATGTGCTGAGGCGAAGATGAAATTTGCAATTAACAACCAAGAGAGAGAAAAC 1500
QY 1454 CACCGGCTATTAACATGCGGATGTGAGATCTGCGTGTTCACAGCAAGAAAGTGAAG 1513
Db 1501 TGTGCAAGTCCGAGAGAAATGATTAAGAGCCAAAGCTGCTTTCAGAGATGCTCAGG 1560
QY 1514 AAGTGTGAGAGGAGAAAGATCAAGAAAGCCCTGCAACCTTGTCTCCGTAATGAAA 1573
Db 1561 AGAATTCAGAGGAGAGAACTCCGAAACATCCCATTAAGTGTGACAGCAAAAGTGA 1620
QY 1574 AAGAGACCTGAGAGGAG-----AAGTAACTCATTTATGCAAGAGGAGAGAGCAG 1624
Db 1621 TCGAGGAGCTGCGAGAGTTGTCTGAGAGTGAAGGGGCTTCAGGCGGACACCGAGAG 1680
QY 1625 GTAGCGTGAAGACCAAGAGACATCAAGGTGACAAAGTTTACAGCGGAGAGAGAA 1684
Db 1681 CACTCGTGAAGACCCGCGGCTCATGTAAGTAATTAACCTCAAGCAAAATGACCGTATGA 1740
QY 1685 TTGGGCTTACGCTATTAATTTCAACCCAGGCGTATTTGAATATGAAAACTGGCGTGA 1744
Db 1741 TCGGACAGTATATGTTGTCTCCGATCTCTGTGTGAAGAACCTTAACTGCAACAG 1800
QY 1745 TCCACCATTTGGCGGAAACAAAGTACTGTGATATGACTCAAAAGGTAGGCGAGGATACA 1804

1801 CACACCCCTGACACACAGGTTAAGATCATACCACATCCCGAAGATCAGAAAGTATG 1860
1805 AAGTGAGCATATACACGGTAAGTCATTGTACGAAAGGAGCGCGGCTCTTCAAG 1864
1861 CAGTGAAACATACACGCTTAAGTATCTGATCCAGCAGAAAGTCCCGTACATGCGCAG 1920
1865 ACTTCAGAGCATTTGAGTGAAGAGCGGTAGCATGTTTCAACGAGAGGAGTTCGTAACA 1924
1921 AATTCTTAGCATGAGTGAAGCGCCAGCTTGTTGACAAAGAAAGAGTTTGGAAAC 1980
1925 GATACCTGACCACTTCGATCAACGAGAGCGCCTAAACCTGACGAAAGTACTATA 1984
1981 GCAACCTGTACCATTTGCACTGACAGCGTCCGCTAAGATACAAAGAGGAGCGTACA 2040
1985 AGACGTAAAGACTCAGGACACAGACTGAGATAGTCTTGATATTGACGACGAAAGT 2044
2041 AGGTTACAAAGGACAGACTCCGAGAAACAGATACGTTGTGACGTGGCAAGAACGAT 2100
2045 GTGTTAAGCAGAAAGCAGAGTCCCTTGCTTAACCGGTATCTGTAAGTCCACAT 2104
2101 GCGTTAAGAGGAAAGAGCTCAGAGACTTGCTTTCGGAGAACTGACCAACCGCCCT 2160
2105 TTCAGAGTTGCGTACGAGAGTCTCAAGACGACGACGACACCTGACAAAGTCCAA 2164
2161 ATCAAGAACTAGCTCTGAGGAGCTGAAGACTGCAACCGCGGTCCCGTACAAAGTTGAAA 2220
2165 CCATCGGAGCTATAGAGATGCGCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGTGA 2224
2221 CAATAGGAGTATAGGACACACAGAGTCCGGCAATCAGTATCAAGTCAACTGCA 2280
2225 CTAAAGAAAGATCTGTTGTGAGTGGAGAAAGAAAGTCCGACAGAAATCATCAGGATG 2284
2281 CGGCACTGATCTTGTATACAGCGGAAAGAAAGAAATCTCGCGCAATTTGAGCGCGACG 2340
2285 TAAAGAGATGAGACGATGATGATGTTGCTGCTAGAGACTGTGATTCAGTCTTCTAAATG 2344
2341 TGCTACGGCTGAGGGGACATGACATCACTGCAAGACAGTGAATTCGGTTATGCTCAACG 2400
2345 GGGTTAACCACCCCTTAACTCTGTACATTTGATGAGGACATTTCCCTGACGAGGA 2404
2401 GATGCCAACAAGCCGCTAGAGATGTTGATGTTGAAGAAAGCGTTCCGATGCCACGAGAG 2460
2405 CGCTGCTGGCACTGATGTCATGCTCAACAC---TAAAGAAAGTGTATTTGTCGGGGAC 2464
2461 CACTACTTGCTGATTTGCAATGCTCAACCCCGTAAGAGTATCTATGCGAGAC 2520
2462 CAATAACAATGCGGCTTTCTTAAATGATGCTGCTGAAGTACATTTTAACT----- 2512
2521 CTAAAGCAATGCGGATTTCTTCAACATGATGCAACTAAAGTACATTTCAACACCGCTGAAA 2580
2513 ATGACATATGCACTGAATGTACATTAAGCACTCTAGAGAGTGCACACAGACTGTA 2572
2581 AAGACATATGTAACCAAGCAATTTCAAGTTATCTCCGACGTGCAACAGCGCACTCA 2640
2573 CCGCATGCTGCTCCACGCTCTTCTAGCAACAGCGAATGAAGCGGTTAAACCATGCTG 2632
2641 CCGCATATGATTCACACTGCACTTTCGATGGAATTAAGAAACCAAAACCGCTGCAAGA 2700
2633 ATAAATATCATATAGATACACAGGAGACACAAAGCGCACAAAGATGATCTGATTTCA 2692
2701 AGAATATGAATGACATTTACAGGGGCGCAGAACCGAAGCGAGGAGCATCATCTCGA 2760
2693 CCTGTTTCAAGAGATGGGTGAACAGCTACAGATTTGACTACAAATAATCAGAAATCATGA 2752
2761 CATGTTTTCGCGGGGTGTTAAGCAATCTGCAATGCAATCTCCGAGATGAGGTAATGA 2820
2753 CTGCGGCTGCACTGCAAGGACTTACGCGGAAAGCGGTTATGCTGCTGAGGTAACAAAGTCA 2812
2821 CAGCGCGGCTTCAAGAGGCTTACCAAGAAAGGATGATGCGCTCGGCAAAAGTCA 2880
2813 ACGAATCACTACTGCAAGCTTCTGAGCAGTGAACGTGTAATCTTACACGCAAG 2872

2881 ATGAACCCGCTGTACCGCATCATCAGAGCATGTGAACGTGTGCTCACCCGACATG 2940
2873 AAAACGACTTGTGCAAGAGCGTACGTGATCCCTGGATTAAGACATTTACAGCTA 2932
2941 AGACAGGCTATATGAAAACTTTACAGGCGACCATGATTTAAGACGCTCACTAACG 3000
2933 AATATCCGGGAGATTTACAGGCTTCAATGGAACGATGCGACGCGCAACAGCAGCATTA 2992
3001 TACCTAAAGAAATTTTCAAGCGCACATCGAGACTGGGAAGCTGAAACAAAGGAATTA 3060
2993 TGGACGCGTCTTGTATAGCCGACAGACGCTGATGTGTTCCAGAAATAGGTAAAGCTCT 3052
3061 TTGCTGCATTAACAGTCCCGCTCCCGTACCAATTCCTTACGCTGCAAGACTAAGCTTT 3120
3053 GCTGGCGAAGGCTTTAGAGCAGCTTGAGCGCGCAACATTTGCTGACGAGACAGC 3112
3121 GCTGGCGAAGCACTGGAACCGATCTGGGCGAGCGCGGTATCTTACTTACCGGTTGCC 3180
3113 AGTGG---GAGACGTTGCAACCATTTCAAGCATGACAGAGCGTACTCACTGAAATGGCAC 3169
3181 AGTGAAGGAGCTGTTCACAGTTTGGATGACAAACACACTCGGCACTTACGCGCT 3240
3170 TGAATCTTTTGAACAGGTTCTTTGGATGAGACTGGAAGTGGTTATTTTCGCTC 3229
3241 TAGAGTAAATTTGCAATTAATTTTTCGCGATGACTTGAACAAGCGGCGCTGTTTCCAAAC 3300
3230 CTACCGTGCACTTATCTTACAGGAT-----CAGCATGGGATA 3268
3301 AGAGATCCGTTAAGTATCATCTCTGCGACTCAGCGAGGCGAGTACCTATGGGACA 3360
3269 ACTGCCGAGGAAACATGATGAGGCTTAATAGAGAGTGAACAAAGAGTTGTACAGCG 3328
3361 ACAGCCAGAAACACAGAGATATGGTATGATCAAGCCGTTGCCGCAACTCTCCGTA 3420
3329 GATATCCGTGATCAAAAACGCGTTGACACAGCGAGGTAGTGTATTAAGAAATATA 3388
3421 GATTTCCGTTGCTCAGCTAC---TGGAAGAGCACACAGCTGATTTGACAGCGGCA 3477
3389 CCATCAAGGACTACTCTCCACATTAATGTGTTCCATTAAATGCGCGTGGCCCACT 3448
3478 GAACTAGATATCTCTGACAGATTAATCTTGTCCTCAGTACCGCAATCTCCCTCACG 3537
3449 CGTTGATGTTGACACAAAGAGCAGGTAACAATGATCAGCGGATTTCTATCTAAGA 3508
3538 CTTATGTCCTCGAGACACAGAGAAACAACCGCGCGGTGAAAATTTCTTGAAGCCGT 3597
3509 TGAAGGCAATCTGTGTGTGATCGCGATCCTA-----TCAGCATTCAGAGG 3558
3598 TCAAAACCACTCGTACTGTGATCTCAGAGAAAAAATGAAAGCTCCCAACAAGAAA 3657
3559 AAGAAATGAGTGCATGGGTCCATTGCCCATAATACATCAGGTGATCTCGATTTG 3618
3658 TCGAATGATGCGCCCAATTTGCAATGCGCGGAGATTAACATCAACCTGCTTTCG 3717
3649 GGAATACCTAGCAATGTCGTAATATGACATTAATCTTTGCAATGTTAAGACCCCGTAC 3678
3718 GCTTTCGCGCGCAGGACG---TAGCATCTGCTTTCATCAATATTTGAACTTAATATC 3773
3679 AGAACAATCACTACCAACAGTGCAGAGATACGCTATCCACACAGCATCTAACGTTG 3738
3774 AGAAACATCACTTTCAACAGTGCAGAGACACAGCGGGACTTGAAGAAACCTTTTCGGGT 3833
3739 AAGGCTGTCCACACTGGAACACTGGCGGAACATGTGTGCTATAGGGTATGGGCTTCT 3798
3834 TCGCGCCCTGAATCTCTTAACCCCGAGGCACTCTGCTGGAAGTCTCTACGTTTACGCC 3893
3799 GATCGCGCAACAGGAATATCATCTACGCGGTGCAAGCTCATTTAAGTTAACCGGTTC 3858
3894 GACCGCAATATGATGAGAGCTGATCAAGCTTTTCCGAAATTTTTCAGAGTGTCTCA 3953
3859 TGTAGCTTAAGAACACTGCGGAATATCTGAGGTTCTTCTGTTCTTTCGGAAGAGAC 3918
3954 GCGAGGCGAGAGTCTGCTCAAGCAATACAGAAATGATACGTGATTTCCGACAACTGAC 4013

Oy	3919	AACGGCAACACACATGATGACAGAGACAGATCCGGGTGATGGCTTGACAAATCATATCA	3978
Db	4014	AACAGCCGCACACAGACAAATTAACCCCGCATCATTTGAAATTTGTGATTTCCGTGTAC	4073
Oy	3979	GGGTCAACCAAGTACGAGGACGAGAGAGCTCCAGCGTACAGATGATCAGAGTGACAT	4038
Db	4074	GAGGTATCAAGAGACGAGATTGGAGCCGACCGGTGTACCGTACTATAAAGGAGAAAT	4133
Oy	4039	AGCAAGACCGTGAACCAAGCTATGCTTATATGCTGTATATGACAAAGTCAACCGCTCC	4098
Db	4134	GCTGATTTGTCAAGAGAAAGACGTTGTCAATGAGCAATCCAATGGGAGACACAGAGAA	4193
Oy	4099	GGAGTGTGCGGTGACCTGTACCGAAATATGGCCGGCTGCTTTTGATAGACACCAATAG	4158
Db	4194	GGAGTGTGCGGTGCATCTATATAACGTTGGCCGAACAGTTTACCGATTCAGCCACAGAG	4255
Oy	4159	GTCCGGACCGCTAGACT--TGTGAAGACGAAACCGCTCATCATCATGCTGTAGAACCC	4215
Db	4254	ACAGGTACCGGAAAATCTGACTGTGTGCCAAAGAAAAGAAATGATGCCAGCGGTGGCCCT	4313
Oy	4216	AATTTTCTAAGATGCGCGGAACCGGAGGGCGACCTTAAGCTCGACGCTGCGCTACATGAGC	4275
Db	4314	GATTTCCGGAAACACCCAGAGGACGAGAGCCCTGAATTTGCTGCAAAAACGCTTACATGCA	4373
Oy	4276	ATPACGTCATCGTCAACGCTGAGCGGATTAACAAAATATCATAGTACCGCTACTGTCAAC	4335
Db	4374	GTGCAGACTTATGTAATTAATGAACATTAATTCAGTGTGCGCATCCACTGTATCTAC	4433
Oy	4336	GGCATCTATTTGTGGGCAAGATTCGATGATGCAATGATTCATCACTGTTCATCTCT	4395
Db	4434	GGCATTTAGCAGCCGGGAAAAGACCGCCTTAGGTATACCTTAAGCTTTGACAAACCGCG	4493
Oy	4396	TTGCACACTACGATGCGATGCGATGTCCATATATTGCTTGATTAACAATGGGAGACGAG	4455
Db	4494	CTAGACAGAACTGATGCGGAGGTAAACCATTACTGCTCGTGATTAAGATGGGAAGAAAG	4555
Oy	4456	ATTAATCGAGGCAATTCACCCGAAAGAACGTC--GAAATTTGTGATGATGACAGCCA	4512
Db	4554	ATCGACGGGTGCTTCAACTTAAGAGTCTGTAACTGACGTGAAGATGAGAGATATGAG	4613
Oy	4513	GTAGACATTTGACTTGTGAGGATCCACCCAAACAGCTCTTTTGGACGAGACAGATTAC	4572
Db	4614	ATCGACGACGATTATGATATGATTCATCCGACAGTGTGCTGGAAGGAAAGAAAGGATTC	4673
Oy	4573	TCCGTCAATGAGGGCAAGTTGATTCATACCTGGAAGGTACACGATTCATCAGACCGCC	4632
Db	4674	AGTACTACAAAGGAAAGTTGATTCGTACTTTGAAAGGACCAAAATTCATCAACAGCACA	4733
Oy	4633	AAGGACATTTGCGGAATTCATGCAATGTGGCCCAACAAATCTGAGCGTAAATGAGAGATT	4692
Db	4734	AAAGTATGCGGAGATTAAGGTCTCTGTTCCAAATGACACAGAAAGCAACGAACTG	4793
Oy	4693	TGCTTGTATCATCTCGGGGAGAGTATGTCCAGCATCCGCTCCAAATGCCCAGTAGAGAG	4752
Db	4794	TGTGCTCATATTGGGGAGACCATGTGAAGCAATCCGCAAAAATATGCCGGTGTGACAC	4855
Oy	4753	TCAGAGGCTGTCTGCCATCTCACCATCTTCATATCCCTGTGTAATTAAGCTATAGCGCT	4812
Db	4854	AACCGTGTCTAGCCCGCCAAAACGCTGCGGTCTCTGTATATGATGCAATGACGCA	4913
Oy	4813	GAGCGGTATACAGTTTGCCTGTGCGAAGAAAGAAAGATTCGCGGTATGCTCATCTTC	4873
Db	4914	GAAAGGTTCACAGATCAGAAAGCAATACGTCAAAAGAAAGTTATCAGTATGCTCTCCAC	4973
Oy	4873	CTGTTGCCAAGTACAGATCACAGGCGTGCAGAAAGTACAGTGCAGCAAAACAGTCTG	4933
Db	4974	CCCCCTCCAAAGTACAAATATCAAGAAATGTTCAAGAGTTCTAGTGCACAAAAGTATGCTG	5033
Oy	4933	TTTTTCAGGCGTGTACCAACCGGCTGTACACCCACAGAAATACCGCGAA	4980
Db	5034	TTTAAACCGCATACCCCGCATTCGTTCCCGCCGCTAAGTAAATATGAA	5081

OY 323 TTAAAGAAA-----GTGACATTACGACAGAAATAGCTCTTAAGGCGGAGACTGEC 376
Db 364 TGGCGAAAAGATGTAAGATTACAAAACAAGAACTTGATGAGAAAGATCAAGAACTCC 423
OY 377 TGGAAATCATGTCAACACGACGAGAGACTTCATCTTGATGTGACACAGACGCCA 436
Db 424 GGACCGTACTGTATCACCGGATGTGAAACCGCATCTCTGCTTCCAAACGATGTTA 483
OY 437 CGTGTAGGTACTTTGGAAGTGTAGCATATACCAAGATGTACCGAGTCCATGACCGA 496
Db 484 CTTGACACACGCGTCCGAGTACTCCGTATGAGAGAGGTGACA---TCAACGCTCCCG 540
OY 497 CATCAATCACCACGAGCGCTTAAAGAGATTAGACATTTACTGATAGGCTTTGACA 556
Db 541 GAACATATTACACACAGGCTATGAAAGGCGTGGACCTGTACTGATGATGCTTCGACA 600
OY 557 GCAGCCCTTTATGTACAAAACATGCGAGGTTCTTACCCTTACTTACAAACAGAACTGGG 616
Db 601 CCACCCAGTTCATGTCTCGCTATGGCAGGTTCCGATCCCTGCATACAAACCAACTGGG 660
OY 617 CTGACGAGAGATTTGGAAGCATGATGAGCATTTGCGCTCGTAACTCAATCTTCAGAGA 676
Db 661 CCGAGAAAAGATCCTTGAAAGCGTAACTGCACTTGACAGCAAAAGCTGAGTGAAG 720
OY 677 GCAGGCTTGAAGAACTCTCAATCCTTAGAGAAAGAGGCTCCAACTTACTTAATAGATCA 736
Db 721 GCAGGACAGAAAGTGTGATATGAGAGAAAGAGTGAAGCCCGGATCAGGGTTT 780
OY 737 TATTTCTGCTGTTTCAACATCTACACAGAAATAGATCACTGTTACGTAGCTGGCATC 796
Db 781 ATTTCTCGGTGATGACACTTTCACAGAACACAGAGCCAGCTTGACAGACTGGCATC 840
OY 797 TTCCAAGGTGTCACCTTGAAAGAAAGTCACTTCAACAGTATGATGTGGACCATTTG 856
Db 841 TTCAATCGGTGTTCACTTGAAAGAAAGCACTGTAACCTTGCGCTGTGATCAAGTGG 900
OY 857 TCAGCTGTGAAGGATGATCATCAAAAGATACATCAGCCAGGACTATACGTTAAAG 916
Db 901 TGAAGTGGAGAGGCTACGTATGTAAGAAATCAACATCAGTCCCGGATCAGCGAGAAA 960
OY 917 TTGAGACTTGGCGTCCACATGATCATCGGAGGTTTCTTGATGTGCAAAAGTCAAGATA 976
Db 961 CCGTGGGATAGCGCGGTATCGTAACATAGGAGGCTTCTTGCTATGCAAAAGTTACGATA 1020
OY 977 CGCTGCGGCGGAGAGGTTTCTTTGCTGTGTATGATGATACAGCCACACTTTGCG 1036
Db 1021 CAGTAAAGGAGAAAGGATATGTTCCCGGTGACATATATCCCGGACACCATATGCG 1080
OY 1037 ATCAGATGACAGGATTTCTGGCACTGACGTTAGTGTGATGAGGACCAAAAGCTATTGG 1096
Db 1081 ATCAATATACCGGCAATATGAGCAGGATATCTCACTGACGATCACAAAACCTTCTGG 1140
OY 1097 TTGGGCTCAACCAAGAGATTGTCTCAATGTGTAGAGACGCAAGAAATATCTAACATATG 1156
Db 1141 TTGGGCTCAACCAAGAGATTGTCTCAATGTGTAGAGACGCAAGAAATATCTAACATATG 1200
OY 1157 AGAATCTATATACCATGTGTGCGCCAGGCGTTTCCAGGTGGCGGTGAACATCGTG 1216
Db 1201 AAAATTTACTTCTGCAATTCATTTGCAAGGTTTACGAAATGCGCCAGAGGCCCAAG 1260
OY 1217 CCGACTTGAAGCAAGAAAGAACTAGGAGGTGCGGAGGCGCACTTTACTATAGGCGTCT 1276
Db 1261 AAGATCTTGACATTAATAAAATGCTGGGCAACAGAGGCGCAAGTTTACTATATGCTCTCT 1320
OY 1277 GCTGGGCTTTTCAAGACCAAGAAATCAATCCATCTACAGAAAGCTGTATGCAAGAA 1336
Db 1321 TGTGGGCTTTTCAAGAAAGTGAACCTGTCTATGCGCCCACTGGAAGCGAGACCA 1380
OY 1337 TTAAGAAAGTACCTGCTTTTGAATCTATTTGATTTCCAGGCTTACAGACCGGAC 1396
Db 1381 TGTAAAGTCCCAAGCTCTTTTATGCGCTTCCCATGTATCCGTATGATGATCTCTTT 1440

OY 1397 TCGATATGCGCTTCCCGCTAGGCTCAAGCTGTCTTGAAACA---ACTGTCAAAACCG 1453
Db 1441 TGCCATGTCCGTGAGGAGAAATGAAATTTGGCATTTCAACCAAGAAAGGAGAAAAC 1500
OY 1454 CACCGCTATTTCAATGCGCGATGTGAGCATCTGCGTGGCTTTACAGCAAGAGCTGAAG 1513
Db 1501 TGCTGCAAGTCCCGAGGAATTAATGATGAGGCCAAGGCTGCTTTCCAGGATGTCTCAG 1560
OY 1514 AAGTGCTGCAAGCGGAAGATTCAGAGAAAGCTGCGACCTTGCTCCCTGAAATATAGAA 1573
Db 1561 AGGAATTCAGAGCGGAGAAAGCTCCGAGAGACCTCCACACTTATGTGCGAGCAAGGTA 1620
OY 1574 AAGACCCGTAGAGCAG-----AAGTAGACTCTATTAATGCAAGGCGAGAGAG 1624
Db 1621 TCGAGGCGAGCTGCGGAATTTGTCTGCAAGTGTGAGAGGGCTCCAGCGGACACCGGAGAG 1680
OY 1625 GTAGCTGAGAGACACCAAGAGACATCAGAGGTGACAAAGTTACCAGCGAAGAGAA 1684
Db 1681 CACTGTGGAAGCCCGCGCGTCAATGAAGATTAATCTTCAAGCAATGACCGTATGA 1740
OY 1685 TTGGGTCTTAAGCTATCTTCACCCCAGGGGTATTAATATGAAAACTGGGCTGTA 1744
Db 1741 TCGACAGTAAATGCTGTCTCGCGATCTGTGTGTAAGAAAGCTTAACTCGACAG 1800
OY 1745 TCCACCCATTTGGCGGAACAACTGTAATGACTCAACAAAGTGTAGGCGAGAGATACA 1804
Db 1801 CACACCGCTAGAGACCAAGTTAAGATCAATACCACTCCGGAAGATCAGGAAGTATG 1860
OY 1805 AAGTGAGCCATACCAAGTAAAGTCAATTTGACAGAAAGGACGCGGTTCCCTGTTCAAG 1864
Db 1861 CAGTCAACCATACGACCGCTAAAGTATGATGACAGAGAAAGTCCGTACCATGCGCAG 1920
OY 1865 ACTTCCAGGCTTATAGTATGAGCGCTTACATCGTTTCAAGAGAGGAGTTCTGTAACA 1924
Db 1921 AATTTTACACTAGTATGAGCGCGCTTGTGTAAAGAAAGAAATTTGTAACC 1980
OY 1925 GATACCTGCAACATCGCATCAACAGGAGGCGCTTAACTGACGAAAGTACTTA 1984
Db 1981 GCAAGCTGTACATATTTGCCATGACGCTCCCGCTAAGAAATACAGAAAGAGAGTACA 2040
OY 1985 AGACTGTAAAGACTCAGACACAGATCTCAAAATACGTTTCAATTTGACGACAGAAAGT 2044
Db 2041 AGGTTACAAAGCAGAGCTCGCAAAACAGAGTACGTGTTGACGTGACAAAGAGCAT 2100
OY 2045 GTGTAAAGCAGAAAGCAGAGCTCCCTGTGCTTAAACCGGTGATGTGATATCAACAT 2104
Db 2101 GCGTTAAGAAAGAAAGCCTCAGAGACTTGTCTTTCGGGAAAGTGAACCAACCGCCCT 2160
OY 2105 TTCAAGAGTTGGCTAGAGAGTCTCAAGACAGCAGCAGCACTCACAAAGTCCGA 2164
Db 2161 ATCAAGAACTAGCTTGAAGGAACTGAAGACTGACCCGCGGTCCGTACAAAGTTGAA 2220
OY 2165 CCATCGAGTCTATGAGATGCGCAGGTTCAAGTAAATCTGAATCATCAAAAGCGCTGTA 2224
Db 2221 CAATAGAGATATAGGACACACAGAGATGCGCAAGTCAATCATCAAGTCACTGTCA 2280
OY 2225 CTAAGAAAGATCTGTTGTGATGTCGAAAGAAAGAAATGCGCAAAATCATCAGGATG 2284
Db 2281 CCGACGTGATCTTGTATACCGCGGAAAGAAAGAAATGCGGAAATTAAGGCGGAG 2340
OY 2285 TAAAGAGATGAGAGATATGATGTGCTGTAGAGCTGTGATTCAGTGTCTTAAATG 2344
Db 2341 TGTACGCTTAAAGGCGATGACATCAGCTGAAAGACAGTGAATTCGTTATGCTCAAG 2400
OY 2345 GGGTTAAGCACCCGTTTAACTCTGTATGATGATGAGCAATTTGCTGCTCAGGAG 2404
Db 2401 GATGCAACAAAGCGGTAAAGTGTATGTTGACAAAGGCTTCGGGTGCAACGAGAG 2460
OY 2405 CGCTGTGAGCACTGATTCCTGCTCAAC---TAAAGAAAGTGTATTTGCGGAGAC 2461
Db 2461 CACTACTTGCCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
OY 2462 CAAAGCAATGCGGCTTCTTAAATGATGTGCTGAAGATGATCAATTTAAAC----- 2512

```
Db 2521 CTAAGCATGGGATTTCTTCAACATGATGCACTAAAGATCATTTCAACACCCCTGAAA 2580
Qy 2513 ATGACATATGCACTGAAGTGTACATTAAGACATCTCTAGAGGTGCACACAGCTTAA 2572
Db 2581 AAGCATATGTACCAAGACATTTCTACAAAGTTATCTCCGAGCTTGGACACAGCCAGTCA 2640
Qy 2573 CCGCATATGCTCCAGCTCTTCTACGACAGGAAAGAAAGGATTAACCATGTGCTG 2632
Db 2641 CGGCTATTTGTATGCACTGCACTTACATGGAAGAAATGAAACCAAAACCCCTGTGCAAGA 2700
Qy 2633 ATAAATATCATATGATATACACAGAGGACACAAAGCCGACAAAGATGATCTGATTTCAA 2692
Db 2701 AGAATATGAAATGTGACATTTACAGGGGCGACAAAGCCGAAACCGAGGGAATCATCTGGA 2760
Qy 2693 CTTGTTTACAGATGGGTGAACAAGCTACAGATTGACTACAAAATTCAGAAATCATGA 2752
Db 2761 CATGTTTCCGGGGGTGGGTTAAGCAACTGCAAAATCGACTATCCGGAATGAGGTTAATGA 2820
Qy 2753 CTGGGGCTGCATTCGCAAGGACTTACGGGAAAGGGGTTATGCTGTACAGGTACAAAGTCA 2812
Db 2821 CAGCCGGGCTCTCAAGAGGCTTAACAGAAAGAGATATATGCGTCCGGCAAAAAGTCA 2880
Qy 2813 ACAGAAATCCACTACTCTGCGACACTTCTGAGCAGTGAAAGCTGTACTTACAGGCAAG 2872
Db 2881 ATGAAACCCGCTGTACCGCATTCATCAGAGCATGTGAACGTGTGCTCACCCGCACTG 2940
Qy 2873 AAAAAGCATGTTGTGGAAGACGCTAGCTGTGATCCCTGATTAAGCACTTACAGTGA 2932
Db 2941 AGGACAGGCTGATGTGAAACCTTTCACAGGGCGAACCCATGATTTAAGCACTTCACTAACG 3000
Qy 2933 AATATCCGGGGGATTTTACGGCTTCATGTGACAGCTGGCAGCGGAAACAGACGCCATTA 2992
Db 3001 TACCTAAAGGAAATTTTCAGGCAACATCGAGAGCTGGAGAGCTGAAACAAAGGAATTA 3060
Qy 2993 TGGCAGCGGTTCTTGATTAAGCCGAGACAGCTGATGTGTTCCAGAAATAGGTGAACGTCT 3052
Db 3061 TTGTGGGATTAACAGTCCCGCTCCCGGTACCAATCCGTTCAAGTCGAAAGATTAACGTTT 3120
Qy 3053 GCTGGGCGAAGGCTTTAGAGCCAGTCTTGCCACGGCCAAACATGTGCTGACAGACAGC 3112
Db 3121 GCTGGGCGAAGGCACTGGAACCGATCTGCGCACGGCCGGATGCTTACTTACCGGTTGCC 3180
Qy 3113 AGTGG---GAGAGCTGGACCCATTTCAAGCATGACAGAGGATCACTCACTGAATAGGAC 3169
Db 3181 AGTGAAGAGAGCTGTTCCACAGTTTGGGAGTGAAGAAACACACTCGGCACTTACAGCTCT 3240
Qy 3170 TGAATCTCTTTGGACCAAGTTCCTTTGGAGTGAAGCTGGACAGTGGGTTATTTTCCGCTC 3229
Db 3241 TAGACGTAAATTTGCACTTAAGTTTTCGGCATGSACTTGAACAAGCGGCTGTTTTCAAAC 3300
Qy 3230 CTACCGTGCACCTTACTTACAGGAT-----CAGCACTGGGATA 3268
Db 3301 AGACATATCCGTTAAAGTAACTCTGCGCATGCGAGGACAGGACAGTACCTATGGGACA 3360
Qy 3269 ACTGGCAGGGAAGAACTGTATGGGCTTAATAGAGAGTAGCAAAAGAGTTGTCACGGC 3328
Db 3361 ACAGCCGGAAGAACGCAAGTATGGGTACGATACAGCCGTTGCGCCGAACCTTCCTGA 3420
Qy 3329 GATATCCGTGATCAAAACGAGTTGACAGAGGAGGTAGCTGATTAAGATTAATA 3388
Db 3421 GATTTCCGCTGTTCCAGCTAGC---TGGGAAAGGACACAGCTTGATTTGGACAGCGGCA 3477
Qy 3389 CCATCAAGGACTACTCTCCAAACATTAATGTGTTCATTAATGCGCGGTGCCCACT 3448
Db 3478 GAACTAGAGTATCTCTCACAAGCATTACTGTCCAGTGAACCGCAATCTCCCTCAG 3537
Qy 3449 CGTTGATGTGTGACCAAGAGACAGGATACACTGATACAGCGGATTCCTATTAAGA 3508
Db 3538 CTTAGTATCCCGAGACAAAGAGAAACAACCGGCGCGGTGAAAATTTCTTGAAGCAGT 3597
Qy 3509 TGAAGGCAAAATCTGTGTGGATCGGCGATCTTA-----TCAGCATTCAGAGG 3558
Db 3598 TCAAACCACTCCGTACTTGATCTCAGAGAAAAAATTTGAAGCTCCCAAGAGAA 3657
Qy 3559 AAGAAATAGTCCATGGGTTCATTTGCCCAATATACATTCAGGTGTATCTGATTTG 3618
Db 3658 TCGATGATATGCGCCGATTTGGCATATAGCCGGCGAGTAAAGACTACACCTGGCTTTCG 3717
Qy 3619 GGAATACCTAGCCATGTGGTAAATATGACATTAATCTTTGTCAATGTTAGAACCCCGTAC 3678
Db 3718 GGTTCGCGCGAGGACCG-----TACAGCTGCTGTGTTCAATATTTGAATTAATAC 3773
Qy 3679 AGGAACCATCACTTACCAACAGTGGAGATTCAGGCTATCCACCAAGCATGCTAACGTGT 3738
Db 3774 AGAAACCATCACTTCAACAGTGGAGACACCGCGCACTTGAAGAACCTTTCCCGT 3833
Qy 3739 AAGCTGTCCACCACTGGAACACTGGCGGAACATGTGTGCTATAGGATTAAGGCTTGTCT 3798
Db 3834 TCGGCTTGAATCTGCTTAAACCCGGAGGACCCCTGTGTGTAAGTCTTACGTTAGCC 3893
Qy 3799 GATCGGCAACCGAATATCATCTGCGGTGGCAGCTCATTTAGGTTTACCCGTTCT 3858
Db 3894 GATCCGAATAGTAGAGAGCTAGTCAACCGCTCTTGCCAGAAATTTGTCAAGTGTCTGCA 3953
Qy 3859 TGTACGCTTAAGAACTGTCCGAAATATCTGAGGTTCTTTCGTTCTTCCGCAAGAC 3918
Db 3954 GCGAGGCCAAGTGTGCTCAAGCAATACAGAAATGTACTGATTTTCCGACAACTAGAC 4013
Qy 3919 AACGGCAACCAACACATGACAGACAGACTGGGTATGCTGTACCAATCTATCAA 3978
Db 4014 AACAGCCGACACAGCAATTTACCCGCAATATTTGAATGTGTGATTTGTCCTGTTAC 4073
Qy 3979 GGGTCAACAGATACGAGGAGAGAGCTTCAAGCTGACAGATGATCAGAGGTGACATT 4038
Db 4074 GAGGTACAAAGAGAGGAGTTGGAGCGGACCGTCTGATACCGTACTTAAAGGAGAACATT 4133
Qy 4039 AGCAAGCGCTGACCAAGCTATCTGTTAATGTGCTAATAGCAAAAGTCAACAGGTTCC 4098
Db 4134 GCTGATTTGTAAGAGGAAGGATGTGTCAATGACGCCATCACTGGCGACAGACAGGAAA 4193
Qy 4099 GGAATGTGGCGTGCATGTACCGAAATGGCCGGCTCTTTGATTAACAGCCAAATAGCT 4158
Db 4194 GGAATGTGGCGTGCATGTATTAACGTTGGCCGAACAGTTTACCGATTAACGCCACAGAG 4253
Qy 4159 GTCCGAGCGGCTAGACT---TGTGAAGCAGAAACCGCTCATCATATGCTGTAGAACCC 4215
Db 4254 ACAGGTACCGGAAACGACTGTGTGCCAAGAAAGAAAGATGATCCAGCGGTTGGCCT 4313
Qy 4216 AATTTTCTAAGATGCCGGAACGGAGGCGACCTTAAGTCCGAGCTGCTTACATGAC 4275
Db 4314 GATTTCCGAAACACCCAGAGGCGAAGGCCCTGAATTTGCTGCAAAACGGCTTACATGCA 4373
Qy 4276 ATAGGTCATCTGTCAACGCTGAGCGGATTAACAAAATATCAAGTACGGCTACTGTCAAC 4335
Db 4374 GTGGCAGACTTAAGTAATGAAATATATCAAGTCTGTGCGCATCCCACTGCTATCTTACA 4433
Qy 4336 GGCATCTATTTCTGGTGCAGAAAGATCGAGTATGATCATTTGATCACTGTTTCACTGCT 4395
Db 4434 GGCATTTACGAGCGGGAAGAACCGGCTTGAAGTATCACTTAATCTGTTGACAAACGGCG 4493
Qy 4396 TTGACACTACGATGCCGATGTCAACATATATTGCTTGATTAACATGGAGACCAAG 4455
Db 4494 CTAGACAGAACTGATGTGGAGAGTAAACATCTACTGCTGTGAATAAGAGTGAAGAAAGA 4553
Qy 4456 ATATACGAGCGCATTTCAACCCGAAAGAAAGGTC---GAAATTCGATGATGACAGCA 4512
Db 4554 ATGACACCGGCTGCTCAACTTAAGAGATCTGTAACTGAAGTGAAGATGAGATGAGAG 4613
Qy 4513 GTAGACATGACTTGTCTCAGGCTCCACCAACAGCTCTTTGGACGCGACAGACGCTTAC 4572
Db 4614 ATGACACAGAGTATGATGATTCATCCGACAGTTGCTGAAAGGAAAGAAAGGATTC 4673
Qy 4573 TCCGTCAATGAGGCAAGTTGTATTTCACTGGAAGGTACAGATTCATCAGACGCC 4632
Db 4674 AGTACTACAAAAGGAAGTTGTATTTGTAAGGACACCAATTCATCAAGACAGCA 4733
```

QY 4633 AAGGACATTCGCGAATCCATGATGAGCCCAAAATCTGAGGCTAATGAGCAGATT 4692
DB 4734 AAGATATGGGGAGATAAAGTCTCTGTTCCAAATGACAGAGAAAGCAAACTAG 4793
QY 4693 TGCCTTATCATCTCTGGGAGAGATGTCACAGCATCCGCTCCAAATGCTCAGTAGAGAG 4752
DB 4794 TGTGCTCATATATTGGGGAGAGACATGGAAGCATCCGGAAAATGCGCGGTGACAC 4853
QY 4753 TCAGAGGGCTGCTGCCACCTTCACACTTCCATGCTGTGTAATTAGCTATGAGGCT 4812
DB 4854 AACCGTGTCTAGCCGCCCAAAAACGCTGCGCTCTGTATGATGATGACATGACGCA 4913
QY 4813 GAGCGCGTATACAGGTTGCGCTCTGCGAAGAAAGAGAGTTGCGGTATGCTCATATTG 4872
DB 4914 GAAAGGCTCACAGACTCAGAAAGCAATACGTCAGAAAGAGTTAGTTGCTCTCCACC 4973
QY 4873 CTGTTGCCGAAATACAGATCAAGGCGTGCAGAAAGTACAGTGCAGCAAAACAGTCTG 4932
DB 4974 CCCCTTCCAAAGTACAAATCAAGATGTTGAGAGGTTGAGTGCACAAAGTAGTCTG 5033
QY 4933 TTTTCAGCGCTGTACACCGGCTGTACACCCCGAAGAGTAGCGGAA 4980
DB 5034 TTTAACCCGCAATCCCGCATTTGTTCCCGCGGTAGTATAGAA 5081

RESULT 10

US-08-801-263A-7

Sequence 7, Application US/08801263A
Patent No. 5811407
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
NUMBER OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 5811407th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,263A
FILING DATE: 19-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-801-263A-7

Query Match 26.7%; Score 1333.2; DB 1; Length 11663;
Best Local Similarity 56.1%; Pred. No. 0;
Matches 2820; Conservative 0; Mismatches 2128; Indels 80; Gaps 13;

QY 23 ATATGAAAGAAATTCACGTTGACTTAGATGCTGACAGCCCGTATGCTCAATGCTTACAGC 82
DB 64 AAGAGCCAGTATGTTACCTAGACGTAGACGACCTCAGAGTCCGTTTGTCTGTGCAACTGCAA 123
QY 83 GAGCGTTTCCAAATTTGAGATCGAAGCAAGGAGGCTCACTGACATGACATGCAATG 142
DB 124 AGAGCTTCCCGCAATTTGAGTAGAGACAGCAGAGGTCACTTCAATATGACATGCTAATG 183
QY 143 CCAAGCGCTTTTGGCATGTGGCAACAAGCTCATTTGAGAGGAAAGTCCAGCCGGGACCAAG 202
DB 184 CCAGAGCAATTTTGGCATGTGGCAAGTAACTGATGAGCTGGAAGTTCTTACACAGCGA 243
QY 203 TTATCTTGAATGGAAGTGAAGTGGCCCGTCAAGATGACATTCGATTCGATCCAGCTCAT 262
DB 244 CGATTTTGAATGAGCAGCGGACCGGCTCGTGAATGTTTTCGAGACCAATGATGCAAT 303
QY 263 GTATCTGCGCTATGATGAAGCGCTGAAGACCCGAGACAGTACAAAGGATGAGCAAGAG 322
DB 304 GCGTTTGGCCCATGCGTATGTCAGAAAGACCCGAGCCGATGATGAAATATGCAAGCAAC 363
QY 323 TTAAGAAA-----GTGACATTAACCGACAGAAATATAGCTCTTAAGCGGAGACCTGC 376
DB 364 TGGCGGAAAAAGCATGTAAAGATTCAAAACAAAGACTTGATGAGAAAGATCAAGAGACTTC 423
QY 377 TGAAGTATGATCAACACAGCAGAGACTCCATCTCTGTATGATGACACAGAGCA 436
DB 424 GAGCCCTACTGATGATACCCGATGCTGAAGCCCATCTCTGCTTCCACACATGTTA 483
QY 437 CGTGTAGTACTTTGGAAGTGTAGCAGTATACCAAGATGTATGACGACAGTCCATGACCGA 496
DB 484 COTGCAACACCGCGTGGCAGTATCCGTCATGACAGAGAGTGAACA--TCAAGCTCCCG 540
QY 497 CATCATATACACACAGCGCTTAAAGAGATTAGACAAATTTATGATGATGAGCTTTGACA 556
DB 541 GAACATATTTACACAGCGCTTAAAGAGCGGCGGACCTGTACTGATGATGATGATGACA 600
QY 557 CGACCCCTTTTATGATCAAAACATGAGGAGGTTCTTACTTACTTAAACAGAACTGAG 616
DB 601 CCACCCAGTATGTTCTCGGCTATGAGAGTTGTTGATCCCTGATACAAACCACTGAG 660
QY 617 CTGACGAGAGATATTGAGACAGCTTAACTATGCGTCTGATCACTCAGATCTTCAGAGA 676
DB 661 CCGACGAAAAAGTCTTGAACGCGGTATCATCGGACTCTGACAGCAAAAGCTGAGTGAAG 720
QY 677 GCAGGCTTGAAAACTCAATCTCTTGAAGAAAGAGGCTCCAACTTAATAAGATCA 736
DB 721 GCAGGACAGAAAGTGTGATTAATGAGAAAGAGAGTGAAGCCCGGCTACCGGTTT 780
QY 737 TATCTGCTGTTGTTCAACATCTTACACAGAAAGTATGATCACTGTTAGTGTGCTGATC 796
DB 781 ATTCTCCGTTGATGATGACACTTTTACCCAGAACACAGGCAAGCTTGAGAGCTGAGATC 840
QY 797 TTCCAAACGTTGTTCACTTGAAGAAAGTCTTAACTTCAAGATGATGAGGAGACATTTG 856
DB 841 TTCCATGAGTTTCCACTTGAAGAAAGAGCTGTATCACTTCCGCTGTATACAGTGG 900
QY 857 TCAGCTGTGAAGGAGTGTATCAAAAGATGATGATGAGCCAGGACTTATAGGTAAG 916
DB 901 TGAGCTGGAAGGCTAGCTAGTGAAGAAATCACCATGATCCGCGGATCAAGGAGAAA 960
QY 917 TTGAAGACTTGGCGTCCCAATGATGCGAGGAGTTCTTGAAGTGAAGTCAAGATA 976
DB 961 CCGTGGATACCGGTTTCAAAACATGACGAGGCTTTCTTATGCAAAAGTACCGATA 1020
QY 977 CGTGGCGGCGAGAGGTTTCTTTGCTGTGTATGATGATGATGATGATGATGATGATGATG 1036
DB 1021 CAGTAAAGAGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1037 ATCAATGACAGGAGATTTGCGAACTGACGTTAGTGTGATGACGACCAAAACTATTGG 1096
DB 1081 ATCAATGACAGGAGATTTGCGAACTGACGTTAGTGTGATGACGACCAAAACTATTGG 1140
QY 1097 TTGGGCTCAACCAAGAGTTGTGTCATGATGATGATGATGATGATGATGATGATGATGATG 1156


```

1141 TTGGGCTCAACAGCAATGATGATTAACGGTAAGCACTAACAGGAACCCAAATTCATGTC 1200
QY 1157 AGAAGTATCTATTAACAGTGTGCGCCAGGCGTTTTCAGGTGGCGCGGTGAACATGTG 1216
Db 1201 AAAATTAATCTTCTGCGCAATATGACAGAGGTTTACCAATGGGCAAGAGGCAAG 1260
QY 1217 CCGACTTGAAGAGAGAAAGAACTAGAGGGTGGCGGAGCGCACTTACTATGAGGCTGCT 1276
Db 1261 AAGATCTTGACATGAAAGAAATGCTGGGACACAGAGAGCGCACTTATATGCTGCT 1320
QY 1277 GCTGGGCTTTCAAGACCCAGAAATCAATCATCTACAGAGACCTGTATCGCAACAA 1336
Db 1321 TGTGGGCTTTGCGACTAAGAAATGCACTGCTTCTATGCGCCACCTGGAACGACAGCA 1380
QY 1337 TTAAAGAAAGTACCTGCGCTTTTGAATCTTATGATTCACGCGCTTACCGACCGGC 1396
Db 1381 TCGTAAAGTCCAGGCTCTTTTATGCGCTTCCCATGTATCGGTATGACTTCTT 1440
QY 1397 TCGATATGGGCTTCCGCGGTAGGCTCAAGCTGCTGTTGAACA---ACTGTCAACCCG 1453
Db 1441 TGCCCATGTGCTGAGGAGAGAGATGAATTTGGCATTAACAACAAAGAGAGAGAAAC 1500
QY 1454 CACCGCTATTACAATGCGCGATGTGAGCATCTGCGTGGCTTACAGCAAGACCTGAG 1513
Db 1501 TGCTGAAGTCCCGAGAGAAATTAAGTATGAGAGCCAGGCTGTTGAGAGATGCTCAGG 1560
QY 1514 AAGTGTGAGAGAGAGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1573
Db 1561 AAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1574 AAGAGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1624
Db 1621 TCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1625 GTAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1694
Db 1681 CACTCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1685 TTGGGTCTTACGCTATCTTCAACCCAGAGGAGATTAAGATTAAGTGAAGAAATGGCGTGA 1744
Db 1741 TCGGACAGTATATGTTGTCTGCGCGATCTCTGTGTGAGAGAGAGAGAGAGAGAGAG 1800
QY 1745 TCCACCATTTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1804
Db 1801 CAGACCGGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1805 AAGTGAAGCAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1864
Db 1861 CAGTGAAGCAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1865 ACTTCCAGGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1924
Db 1921 AATTTCTTAGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
QY 1925 GATPACTGCAACCATGCAATCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1984
Db 1981 GCAAGCTGTACATTAATGCGATGCGGTCCCGCTAAGAAATCAAGAGAGAGAGAGAG 2040
QY 1985 AGACTGTAAAGACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2044
Db 2041 AGGTTAACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
QY 2045 GTGTAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2104
Db 2101 GCGTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 2105 TTCAAGAGTTTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2164
Db 2161 ATCAAGAGTACTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2165 CCATGAGAGTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2224

```

```

2221 CAATAGAGTGAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY 2225 CTAAAGAAAGATCTGTTGTGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2284
Db 2281 CCGCAGTGAATCTTGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
QY 2285 TAAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2344
Db 2341 TGTACAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
QY 2345 GGGTTAAGACCCCGTTAACAATCTGTACAATTGATGAGAGAGAGAGAGAGAGAGAG 2404
Db 2401 GATGCGCAAGAGCGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
QY 2405 CCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2464
Db 2461 CACTACTTGTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
QY 2462 CAAAGCAATGCGGCTTCTTAAACATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2512
Db 2521 CTAAAGCAATGCGGATTTCTTCAATGATGCACTAAAGAGAGAGAGAGAGAGAGAGAG 2580
QY 2513 ATGACATATGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2572
Db 2581 AAGCATATGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
QY 2573 CCGCATATGCTTCCAGGCTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2632
Db 2641 CCGCTATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
QY 2633 ATAAATCATATAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2692
Db 2701 AGAAGATGAAATGACATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
QY 2693 CTTGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2752
Db 2761 CATGTTTCCGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
QY 2753 CTGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2812
Db 2821 CAGCGCGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
QY 2813 ACGAGATTCATCTACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2872
Db 2881 ATGAAACCCGCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
QY 2873 AAAAGCATTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2932
Db 2941 AGGACAGCTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
QY 2933 AATATCCCGGGGATTTGACGAGCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2992
Db 3001 TACTTAAGAGAAATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
QY 2993 TGGCAGCGTCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3052
Db 3061 TTGCTGAGATTAACAGTCCGCTCCGCTACATCCGTTCAAGTGAAGAGAGAGAGAGAG 3120
QY 3053 GCTGGCGGAAGGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3112
Db 3121 GCTGGCGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
QY 3113 AGTGG---GAGAGGTTGACCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3169
Db 3181 AGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
QY 3170 TGAATCTTTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3229
Db 3241 TAGAGATTAATTTGATTAAGTTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
QY 3230 CTACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3268
Db 3301 AGAGCATCCGTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360

```


QY 3269 ACTCCAGGGAAGACATGTATGGCTTAAATAGAGGTAGCAAGAGTTGTACGGC 3328
 DB 3361 ACAGCCAGGAACAGCAAGTATGGGTACGATCAGCCGTTGCCCGCAACTCTCCGTA 3420
 QY 3329 GATATCCGTGCATCAAAAAGCGTTGACACAGGAGGTAGCTGATTAAGAAATA 3388
 DB 3421 GATTTCCGGTGTTCAGCTAGC---TGGAAAGGACACAGCTTAATTGACAGGGGA 3477
 QY 3389 CCATCAAGAGACTACTCTCCAAATTAATGTGTTCAATTAATGGCGGTGGCCCACT 3448
 DB 3478 GAACTAGATTATCTTGTGACAGCATTAATGTGTCAGTGAACGCAATCTCCCTCAG 3537
 QY 3449 GGTGATGTTGACCAAAAGGACAGGATCAATGATACAGCGGATCTATCTAAGA 3508
 DB 3538 CTTTATGCCGAGACAAAGAAACAACCGGCCGCTGCAAAAATTTCTTGACCACT 3597
 QY 3509 TGAAGGCAATCTGTGTGTGATCGCGGATCTA-----TCAGATTCAGGG 3558
 DB 3598 TCAACACCACTCCGTACTTGTGATCTCAGAAAAAATTGAAGCTCCCAAGAGAA 3657
 QY 3559 AAGAAAGTATGATCGATGGGTCAATGGCCACTAATACATAGGTGATTCGATTTG 3618
 DB 3658 TCGAATGATGCCCCGATTGGCATAGCCGCGCAGATTAAGAACTAACACTGGCTTCG 3717
 QY 3619 GGAATACCTAGCATGTGCTAATATGACATTTCTTTGCAATGTTAGACCCCGTAC 3678
 DB 3718 GGTTCGCCCGCAGGACGG---TACGACTGTGTTTCAATTAATTGAACTAAATAC 3773
 QY 3679 AGGAACCATCACTCAACAGTGCAGAGATCAAGCTATCACACAGCATGCTAACGT 3738
 DB 3774 AGAAACCATCACTTCAACAGTGCAGAGATCAAGCGGAGCACTTGAACCTTCGGGT 3833
 QY 3739 AAGGTCTCCACCTGAACATGTCGGGAAATGTGTGCTAATAGGTTATGGGCTTCT 3798
 DB 3834 TCGGCTCTGAATGCTTTAAACCCGAGGAGACCTGTGTGTAACTCTCAAGCTTACCC 3893
 QY 3799 GATCGCGAACCAGAAATATCATCTGCGGTGACGCTCATTTAGTTTACCCGTGTC 3858
 DB 3894 GACCCCAATAGTGAAGAGTATGATCACCGCTTTCGCAAAAATTTGTCAAGATGCTGCA 3953
 QY 3859 TGTCACTTAAGAACATGTCGCCAAATACTGAGGTTCTTCTGTTCTTTCGCAAGAC 3918
 DB 3954 GCGAGGCGCAGAGTGTCTCAAGCAATACAGAAATGATACCTGATTTTCGCAACTAGAC 4013
 QY 3919 AAGCGCAACCAACATGACCAAGACAGACTCGGTGATGCTTGAACAATCTATCAA 3978
 DB 4014 AACAGCCGACACGCAATTCACCCCGCATTTGAATGTGTATTTGCTCGCTGAC 4073
 QY 3979 GGGTCAACAGGTACGAGGACAGGAGCTCCAGGTTACAGAGTATCAGAGGTACATT 4038
 DB 4074 GAGGTACAAAGAGCGAGTTGAGCCGACCGTGTACCTTAACCTTAAGGAGAACTT 4133
 QY 4039 AGCAAGACCGCTGACCAAGCTATCTGTTATGCTGCTAATAGCAAGGTCAACAGATTCC 4098
 DB 4134 GCTGATGTCAAGAGAGAGCTGTGATGACGCAATTCACCTGGGACCAAGAGAA 4193
 QY 4099 GGAATGTCGGTGCATGTAACGAAATGCGCGCTGTTTGTATGACAGCCATAGCT 4158
 DB 4194 GGAATGTCGGTGCATGTAACGAAATGCGCGCTGTTTGTATGACAGCCATAGCT 4253
 QY 4159 GTCCGAGCGGTAGACT---TGTGAAGACGAAACGCTCATCATGCTGTAGAGCC 4215
 DB 4254 ACAGGTACCGCAAACTGACTGTGTGCAAGAAAGAAATGATACAGCGGTGGCTT 4313
 QY 4216 AATTTTCTAAGATGCGGAAACGAGGCGCACTTAAGCTGCGAGCTGCTACATGAGC 4275
 DB 4314 GATTTCCGGAACACCCAGAGGACAGAGCCCTGAATTTGCTCAAAAGCCCTACATGCA 4373
 QY 4276 ATAGGTGATCTGTAACGCTGAGGAGATTAACAAATATCATGTAACGCTACTGTCAAC 4335
 DB 4374 GTGGAGACTTATGTAATGAACATTAATCATGCTGTGCGCATCCACTGCTATCTACA 4433

QY 4336 GGCATCTATCTGTGTGCAAGATGAGTATGATCAATTCATCACTGTTTCACTGCT 4395
 DB 4434 GGCATTTACGAGCGGAAAGAACCGCTGTGAGTATCACTTAATCGCTTGACAAACCGG 4493
 QY 4396 TTGCACCTAGGATGCGGATGTCACCATTAATTTGCTTGATTAACATTTGGAGACGAG 4455
 DB 4494 CTAGACAAATGATGCGGACGTAACCATTAATCTGCTGAGTAAAGTGAAGAAAGA 4553
 QY 4456 ATTAATGAGGCAATTCACCGCAAGAAAGCGTC---GAAATTCGATGATGACAAGCA 4512
 DB 4554 ATCGACGCGGTGCTCAACTTAAGAGTGTGTAATCTGAGCTGAAGGATGAGATTTGAG 4613
 QY 4513 GTAGACATTAATGCTGTCAGGATCCACCAACAGCTTTTGGCAGGACAGACAGATTAC 4572
 DB 4614 ATGACAGAGATTATGATGATCAATCGGACATTCCTGAAGGAAAGAAAGGATTC 4673
 QY 4573 TCCGTAATGAGGCAAGTTGATTAATCACTGGAAGTACACGATTCATTAAGACCGC 4632
 DB 4674 AGTACTACAAAGAAAGGATGATTCGTAATCTTGAAGGACACCAATTCATCAACAGCA 4733
 QY 4633 AAGGACATTCGCAATTCATGCAATGTCGCCCAACCAATCTGAGGCTTAATGAGAGAT 4692
 DB 4734 AAGATATGCGGAGATTAAGGTCTGTTCCAAATGACAGAGAAAGCAAGCAACTG 4793
 QY 4693 TGCTTGATCATCTGCGGAGAGATGATGTCAGACATCCGCTCCAAATGCCCAGTAGAGAG 4752
 DB 4794 TGTGCTTACATTAATGGGGGAGACCATGGAAGCAATCCCGAAATATGCCGTTGACAC 4853
 QY 4753 TCAGAGGCTGTGCTCACTTCAACACTTCATGCTGTGTATTAATGCTATGACGCT 4812
 DB 4854 AACCGGTGCTAGCCCGCAAAAAGCTGCGTGTCTGTATGATAGCATGACGCA 4913
 QY 4813 GAGCGCTTACAGTTGCGCTCTCGCAAGAAAGAAAGAGTTCGCCGTATGCTCATATTC 4872
 DB 4914 GAAAGGTTCACAGCTCAGAAAGCAATACGTCAAAAGATTAAGTATGCTCTCCACC 4973
 QY 4873 CTGTTGCCGAATCAGATCAACAGCGCTGCAAGAGTACAGTGCAGCAAAACCACTCTG 4932
 DB 4974 CCGCTTCAAAATCAAAATCAAGATTTCAAGAGTTCAAGAGTTCACTGACCAAAATGATCTG 5033
 QY 4933 TTTTCAGGCGTGTACACCGGCTGTACACCCCGAGAAATGACGCGAA 4980
 DB 5034 TTAAACCGCATACCCCGCATGCTGTTCCCGCGTAGTACATGAA 5081

RESULT 11
 US-09-102-248-1
 ; Sequence 1, Application US/09102248
 ; Patent No. 6008035
 ; GENERAL INFORMATION:
 ; APPLICANT: Johnston, Robert E.
 ; APPLICANT: Davis, Nancy L.
 ; APPLICANT: Simpson, Dennis A.
 ; TITLE OF INVENTION: System for the In Vivo Delivery and
 ; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Bell Selitser Park & Gibson, P.A.
 ; STREET: 1211 East Morehead Street
 ; CITY: Charlotte
 ; STATE: No. 6008035th Carolina
 ; COUNTRY: USA
 ; ZIP: 28234
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: US/09/102,248
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 60..7559
FEATURE:
NAME/KEY: CDS
LOCATION: 7608..11342
US-09-102-248-1

Query Match 26.7%; Score 1333.2; DB 3; Length 11663;
Best Local Similarity 56.1%; Pred. No. 0;
Matches 2820; Conservative 0; Mismatches 2128; Indels 80; Gaps 13;

23 ATATGAAAGATTCACGTTGACTTGAATGCTGAACGCCGTATGTCAGTCTTACAC 82
Db AGAAGCCAGTAGTTACGTAGACCTCCTCAGAGTCCCTTGTGTGCACTGCAAA 123
Qy GAGCGTTTCCACATTTGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 142
Db AGAAGCTTCCGCAATTTGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 124
Qy 124 AGAAGCTTCCGCAATTTGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 183
Db 143 CCAAGCGTTTCCGCAATTTGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 202
Qy 184 CCAAGCGTTTCCGCAATTTGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 243
Db 203 TTAATCTTGAACATTTGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 262
Qy 244 CGATTTTGGACATTTGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 303
Db 263 GTATCTGCGCCATATGTAAGGCGTGAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAG 322
Qy 304 GCGTTTCCGCAATTTGAGATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 363
Db 323 TTAAGAAA-----GTGACATTTACCGACAGAAATAGCTTGAAGCGGAGGAGGAG 376
Qy 364 TGGGCGAAAAGCATGTAAGTTTCAAAAGAACTTGCATGAGAAGATCAAGAGCTCC 423
Db 377 TGAAGTCAATGTAACACGAGCGCAAGACTCTCATCTTGTGTATGACACAGAGCGCA 436
Qy 424 GGAACCGTACTTGAATACACCGGATCTGAAGCCATCACTCTGCTTCCACACATGTTA 483
Db 437 CGGTAGATCTTGAAGTGTACAGATATACAAAGATGTGTAGGAGCTCATCAACCGA 496
Qy 484 CTGCAACAGCGGTGCGAGTACTCCGTATGAGAGAGGTGTA---TCAACGCTCCG 540
Db 497 CATCAATCTACCAACGAGCGCTTAAAGAGTTAGAACATTTACTGTAGTAGCTTTGACA 556
Db 541 GAATCTATTTACCAACGAGCTTAAAGCGGTGCGAGCCCTGTACTGATGGCTTTCACA 600
Qy 557 CGACCCCTTTTATGTACAAAACATGCGGTTCTTACCTTACTTACCAACGAACTGGG 616
Db 601 CCAACCCAGTTTATCTTCTCGGCTATGCGAGGTGTGTAACCTGTACCAACCACTGGG 660
Qy 617 CTGACGAGAGATTTGGAAGCACTTAATTTGGCTTCGTGAATCTCAGATTTTACAGAGA 676
Db 661 CCGAGCAAAAGATCTTGAAGCGGTAACTGGAATCTTGAAGCAAAAGTTGAGTAG 720
Qy 677 GCAAGCTTGAAGAACTCTCAATCTTAGAAGAAAGGCTCAACCTACTAATAGATCA 736

Db 721 GCAAGACAGAAAGTTGCTCATATATGAGAAAGAGGTTGAAGCCCGGCTACCGGTTT 780
Qy 737 TATTCGGTGTGTTGTAACAATCTACACAGAAATGATACATGTTACGTAGCGGATC 796
Db 781 ATTTCTCGTTGATGCACTTTTACCAAGACAGAGCAGCTTGACAGCTGGGATC 840
Qy 797 TTCAAAGCTGTTTCACTTAAAGAAAGCTTAATCTTACAGAGTGAATGTGGACCATTTG 856
Db 841 TTCCATGGGTGTTTCACTTAAAGAAAGAGAGTGTACATTTGGCGGTGTGATACATGG 900
Qy 857 TCAGCTGTGAAGGATCATCAAAAAGATTAACGATCAGCCCAAGACATTAAGGTAAG 916
Db 901 TGAGCTCGAAGGCTACGTAGTGAAGAAATACATCAGTCCCGGATCACGGAGAAA 960
Qy 917 TTGAGAACTTGGGCTCCACATATGATCGCAGAGTTTCTTGAATGTAAGTCAAGATA 976
Db 961 CCGTGGGATACGGGTTTAAACAAATTAAGCAGAGGCTTCTGTATGCAAAAGTTACGATA 1020
Qy 977 CGCTGCGGCGAGAGGAGTTTCTTTTGTGTGTATGATGATGATGATGATGATGATGATG 1036
Db 1021 CAGTAAAGAGAGACGGGATATCGTTCCCGGTGTGACGTATATCCCGCCACCATATGCG 1080
Qy 1037 ATCAGATGACAGGATTTTGGCAATGACGTTAGTGTATGACGACAAATTAATTTGG 1096
Db 1081 ATCAGATGACCGGATATATGACGATATATCTCACTGACGATGACAAATTAATTTGG 1140
Qy 1097 TTGGGCTTCAACCAAGATTTGTGTATGATGATGATGATGATGATGATGATGATGATG 1156
Db 1141 TTGGGCTTCAACCAAGATTTGTGTATGATGATGATGATGATGATGATGATGATGATG 1200
Qy 1157 AGAATCTATTAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1216
Db 1201 AAAATTAATCTTGTGCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Qy 1217 CCGACTTGAACAGAGAAAGAACTAGAGGATGCGGAGGCACTTTACTATGAGGCTGT 1276
Db 1261 AAGATCTTGAACAGAGAAAGAAATGCTGAGGCAACAGAGGCAAGCTTAATATGAGCTGT 1320
Qy 1277 GCTGGGCTTCAACAGACCAAGAAATATCATCTTCAAGAAAGCTGTATGATGATGATGAT 1336
Db 1321 TGTGGGCTTCAACAGACCAAGAAATATCATCTTCAAGAAAGCTGTATGATGATGATGAT 1380
Qy 1337 TTAAGAAAGTACCTGCGCTTTTGAATGATTTTCAAGGCTTATCAAGGAGGAGGAGGAG 1396
Db 1381 TCGTAAAGTCCAGGCTCTTTTGAAGGCTTATGATGATGATGATGATGATGATGATGATG 1440
Qy 1397 TCGATATGAGGCTTCCGCTGAGGCTCAAGCTGTGCTTGAACCA--ACTGTCAACCCG 1453
Db 1441 TGCCATGTGCTGAGGAGCAAGATGAAATTTGCAATTAACCAAGAGAGAGGAGAAAC 1500
Qy 1454 CACCGCTATTAATCAATGCGGAGTGTGAGCATCTGCGGTGCTTACAGCAAGAAAGCTGAG 1513
Db 1501 TGCTGCAAGTCCCGAGAGAAATGATTAAGAGGCAAGGCTGCTTTCAGAGAGTCTGAGG 1560
Qy 1514 AAGTGTGAGGAGGAGAGATCAAGAGACCTTCCAGCTTGTCTCTGAAATAGAAA 1573
Db 1561 AGGATCAAGAGGAGAGAGCTCGAGAGACATCTCCACATTAATGAGCAGACAAAGTGA 1620
Qy 1574 AAGAGCCGTAGAGGAG-----AAGTGAACCTCAATTAATGCAAGAGGAGAGGAGGAG 1624
Db 1621 TCGAGGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
Qy 1625 GTAGCGTGAAGACACACAGAGACATCTGAGGTGACAAATTTACCCAGGAGGAGAGAGA 1684
Db 1681 CACTGTGCAAAACCCCGCGGTCTATGTAAGATTAACCTCAAGCAAAATGACCGTATGA 1740
Qy 1685 TTGGGCTTATGAGTATCTTCAACCCAGGAGGATTAATGAGAAATCTGGCGTGA 1744
Db 1741 TCGGACATTAATCTGTTGTTCTTCCGCAATCTTGTGCTGAAGAGCCTAACTGCAACAG 1800
Qy 1745 TCGACCATTTGGGAGCAAGATCTGATTAATGATCAAAAGTGGAGGAGGAGATACA 1804

Db 1801 CACACCCGCTAGACAGACAGGTTAAGATCAATGACCTCCGAAAGATCAGAAAGATATG 1860
QY 1805 AAGTCAGACATATACACGCGTAAAGTCAATTTGTAACAAGAGGAGCGGCGGCTCGTTCAAG 1864
Db 1861 CAGTGGAAACATATACAGCGCTAAAGTATATGTCACAGAGAAAGTCCGTTACCAATGGCCAG 1920
QY 1865 ACTTCAGAGCAATYAGTAGAGACCGCTACGATCGTTTCAACGAGAGGAGTTCTGTAACA 1924
Db 1921 AATTCTTAGCACTGAGTAGAGCGCCAGCTTGTGTACAAACAAAGAGAGTTTGTGAACC 1980
QY 1925 GATACCTGACCACTATCCGATCAACAGAGAGCCGTAACCTGACGAAAGTACTATA 1984
Db 1981 GCAACCTGATCCATATTTGCACTGACGCGTCCGCTAAGATACAGAAAGAGAGCACTACA 2040
QY 1985 AGACTGTAAAGCTCAGAGCAGACAGACTGAAATAGGCTTGATATGACGACGAAAGT 2044
Db 2041 AAGTTACAAAGGCAAGAGCTCCGCAAAAAGAGTATGTTTGAAGCTGGCAAGAAAGCAT 2100
QY 2045 GTGTTAAGCGAAGAGACGAGGCTCCCTGTGCTTAACCGGATCTGTAGATCCACCAT 2104
Db 2101 GCGTTAAGAGAGAAAGAGCTCAGAGACTTGTCTTTCCGAGAACTGACCAACCCGCCCT 2160
QY 2105 TTCAAGAGTTTCCGTAACAGAGTCTCAGACAGACAGACAGACCTCAAGAAATGCCAA 2164
Db 2161 ATCAAGAACTAGCTTTGAGGAGCTGAAGACTGACGACCCGCTCCGTAAGAGTTGAAA 2220
QY 2165 CCATCGAGTCTATGAGAGTCCAGGTTCAAGTAATCTGGAATCAATCAAAAGCGCTGTA 2224
Db 2221 CAATGAGAGTATAGACACACAGAGTCCGCAAGTCACTATCAATCAAGTCAACTGTA 2280
QY 2225 CTAAAGAAAGATCTGTTGTGAGTGCAGAAAGAAAGAACTCCGCAAGATCATCAGGAGT 2284
Db 2281 CGGCAAGATCTTTGTAACACAGGAAAGAAAGAAAGAACTCCGCAAGATCATCAGGAGT 2340
QY 2285 TAAAGAGATGAGACGATGATGTTGCTGCTAGAGACTGTCAATTCAGTCTTCAATG 2344
Db 2341 TGCTACGCTGAGGAGCATGACATCACGTCGAAAGACAGTGAATTCGTTATGCTCAAG 2400
QY 2345 GAGTTAAGACCCCGTTAAACCTCTGTACTTGAAGAGGCAATTCGCTCCGATGAGAGGA 2404
Db 2401 GATGCCAAGAGCCGTTAGAGTGTGTATGTTGAGAGAGCGTTCCGCTCCAGAGAGG 2460
QY 2405 CGCTGCTGAGCACTGATGCTCATGCTCAAAAC---TAAAGAAAGTGTATTTGTCGGGAGC 2461
Db 2461 CACTACTTGCTGTATGCAATGCTCAGACCCCGTAAGAAAGTATGATCTATGCGAGAC 2520
QY 2462 CAAAACAAATGCGGCTTTCTTAAATGATGCTGCTGAAGTACATTTTAAAC----- 2512
Db 2521 CTAAACAAATGCGGATTTCTTAAATGATGCACTAAAGGTACATTTCAACACCTGAAA 2580
QY 2523 ATGACATATGACCTGAAGTATCCATTAAGACATCTTAGAGAGTGCACAGACTGTAA 2572
Db 2581 AAGACATATGTAACAAGACATTTCTACAGTTTATCTCCGACGTTGCAACAGGACGTA 2640
QY 2573 CCGCATGCTGTCAACGCTCTTCTACGACAAAGCAGATGAAGAGCGGTATTCATGCTG 2632
Db 2641 CCGCATATGTAATGACACTGCAATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2700
QY 2633 ATAAATCATATAGATATCACAGAGGACCAAGAGCCGCAAGAGATGATCTGATTTAA 2692
Db 2701 AAGACATGAAATCGACATTAACAGAGGACCAAGAGCCGCAAGAGATGATCTGTA 2760
QY 2693 CCGTTTGAAGAGAGGAGGAGAAACAGCTACAGATTAAGTGAAGTGAAGTGAAGTGA 2752
Db 2761 CATGTTCCGCGGAGGAGTGAAGCACTGCAAAATGATCTCCGACATGAGTGAAGTGA 2820
QY 2753 CTGCGGCTGCAATGCAAGAGCTTACGCGGAAAGCGTTATCTGTCAAGTCAAGTCA 2812
Db 2821 CAGCGCGGCGCTCAACAAGGCTTAACCAAGAAAGAGTATATGCGCTCCGCGCAAAAGTCA 2880
QY 2813 ACAGAGATCACTTATCTGACAGACTTGTAGACAGTGAAGTGTATCTTACAGGACAG 2872
Db 2881 ATGAAACCCGCTGATCGGATCATCATGACAGATGTGAAGTGTGCTCAACCCGACCTG 2940

QY 2873 AAAACGATTTGTCTGAAAGACGCTAGCTGTGATCCCTGATTAAGACATTTACAGCTA 2932
Db 2941 AGGACAGGCTATATGAAAACTTTACAGGGGACCCATGATTTAAGACCTCATTAAG 3000
QY 2933 AATATCCCGGAGATTTACAGGCTTATGACGATGCGAGGCGCAACGACGCGCATTA 2992
Db 3001 TACTTAAAGAAATTTTACAGGCAACATCGAGAGCTGGGAGCTGAAACCAAGGAAATA 3060
QY 2993 TGGCAGGCTTTTATTAAGCCGCAAGACTGATGTGTTCCAGATTAAGTGAAGTCT 3052
Db 3061 TTGCTGCAATTAACAGTCCCGCTCCCGTACCAATCCGTTGCTGCAAGACTTAAGT 3120
QY 3053 GCTGGGCGAAGCTTTAGAGCAATGAGCAATGAGGCGCAACATTTGCTGAGAGACAG 3112
Db 3121 GCTGGGCGAAGCACTGAAACGATATCTGACAGGCGGATATGTTACTTACCGTTGCT 3180
QY 3113 AGTGG---GAGAGCTTGACCCCATTCAGCATGACAGAGCGTACTACCTGAATGGCAC 3169
Db 3181 AGTGAAGGAGGCTGTTCCACAGTTTGGGAGTGAACAAACCACTCGGCTCATAGCCT 3240
QY 3170 TGAATCTTTTGGACAGGTTCTTTGAGTGAACCTGACAGTGGGTTATTTCCGCTC 3229
Db 3241 TAGAGTAAATTTGCAATTAAGTTTTCGCAATGACTGACAGCGGCTGTTTCCAAAC 3300
QY 3230 CTACCGTGCATTTACTTAACAGGAT-----CAGCACTGGGATA 3268
Db 3301 AAGCATCCCTTTAAGTACATCTCTGCCACTCAGAGGCGCAAGTACTGCTTGGGACA 3360
QY 3269 ACTCCAGGAGAAACATGATGGCTTAAATGAGAGTACAAAGAGTGTGACAGCG 3328
Db 3361 ACAGGCTAGGAAACAGCAAGATATGATGATGATGATGATGATGATGATGATGATGAT 3420
QY 3329 GATATCCGTGATCAAAAGCGGTTGACACAGGAGGATGATATTAAGAAATATA 3388
Db 3421 GATTTCCGCTGTTCCAGTACG---TGGGAAAGGACACAGCTTATTTGACAGAGGCA 3477
QY 3389 CCATCAAGAGCTACTCTCCAAATTAATGAGTGTCAATTAATGAGGAGTGTGAGGAGT 3448
Db 3478 GAACTAGATATCTTGTGACAGATTAATGATGATGATGATGATGATGATGATGATGAT 3537
QY 3449 CGTTGATGTTGACCAACAAAGGACAGGATACAACTGATCAACAGGATTTCTATTAAGA 3508
Db 3538 CTTATGCTCCGAGAGCAAGAGGAAACAAACCGGCGCGGTGGAATAATCTTGAACGAGT 3597
QY 3509 TGAAGGCAATCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3558
Db 3598 TCAAAACACACTCCCTACTTGTGATCTGAGAGAAATAATGAAAGTCCCAACAGAGAA 3657
QY 3559 AAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3618
Db 3658 TCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3717
QY 3619 GGAATACCTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3678
Db 3718 GGTTCCTCCGCGAGCAAGG---TAGACCTGAGTGTCTATCAATATGGAATGTAATTA 3773
QY 3679 AAGAACTATCTACCAACAGTGGAGATGACGCTATCAACACAGCATGCTTAAGTGT 3738
Db 3774 AAGAACTATCTATCTTCAACAGGCGCAACAGCAGGCGGACCTTGAACCTTTGCGCT 3833
QY 3739 AAGGCTGTCACCACTGAAACATGAGGAGGAAACATGAGGAGTGAAGGATGAGGAGTGT 3798
Db 3834 TCGGCTCTGAATCTTAAACCGGAGGACCTCGTGTGAGTGTCTTACGTTTACGCT 3893
QY 3799 GATCGCGCAACGAAATATCATCTGCGGTGACGCTCATTTAAGTTTACCGGTGT 3858
Db 3894 GACGCAATATGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3953
QY 3859 TGTACGCTTAAGAACATGCGGAAATATGAGGTTCTTGTGTTGTTCTTGGCAGAGAC 3918
Db 3954 GCGAGGCGCAGATGCTGCTCAAGCAATACAAATATGATCTGATTTTCCGACCACTAGAC 4013

```

3919 AACGGACACACACATGACGAGACAGACTCGGTGATGCTTGAACAATCTATCAA 3978
4014 AACGCGGCACACAGACATTCACCCCGCATCTTGAATTTGTGATTTCTCCGTATC 4073
3979 GGGTCAACCAAGTACGAGGACGAGAGCTTCACCGTACAGATGATCAGAGTGACATT 4038
4074 GAGGGTCAAGAGACGAGAGTGGAGCCGACCGTGTATCCGTAATAAAGGAGAACATT 4133
4039 AGCAAGAGCGGTGACCAAGCTATGCTTAATGCTGTAATAGCAAGGACGAGTTCC 4098
4134 GCTGATTTGCAAGGAGAGAGGAGTGTCAATGACCAATTCACCTGGGACGACGAGAA 4193
4099 GGAAGTGGCGGTGACGTAACCGAAATGCGCGCTCTTTGATAGACAGCAATAGCT 4158
4194 GGAAGTGGCGGTGACGTAACCGAAATGCGCGCTCTTTGATAGACAGCAATAGCT 4253
4159 GTCGGACGCGTACACT--TGTGAGACACGACCGCTCATATACATGCTGTAGAACCC 4215
4254 ACAGGTACCCGAAATGACCTGTGTGCAAGAAAGAGATCCACGCGGTTGGCCCT 4313
4216 AATTTTCTAAGATGCGGAAACCGGAGGCGACCTTAAAGTCCGAGCTGCTATAGAC 4275
4314 GATTTCCGAAACACCCAGAGGACGAAAGCCCTGAATTTGTCGAAACGCTTACATGCA 4373
4276 ATAGCGTCAATGTCACAGCTGACGAGATTCAAAATATCATACCGCTACTGTCAACC 4335
4374 GTGGCAAGCTTAAATGAAATCAATATCAAGTCTGTGCGCATCCAGCTGCTATCA 4433
4336 GGCATCTAATCTGTGTCGAAAGATCGAGTGAATGCAATCATGTCACCTGTTCACTGCT 4395
4434 GGCATCTAATCTGTGTCGAAAGATCGAGTGAATGCAATCATGTCACCTGTTCACTGCT 4453
4396 TTGACACTACGAGTTCGCAATGTCACATATATGCTGTGATAAACAATGGGAGAACAGG 4455
4494 CTAGACGAATCTGATGCGAGATGACATCTGCTGATGAGAAAGAGGAGAAAGAA 4553
4456 ATATCAGAGCCATTCACCGCAAGAAAGACGCT--GAAATTCGATGATGACAGCA 4512
4554 ATGACCGCGGTGCTCCAACTTAAAGAGCTGTAACTGAGAGATGAGATGAGATGAG 4613
4513 GTAGACATTAATCTGTGTCAGAGGTCACCCAAACAGCTTTTGGAGGACAGCAGTTAC 4572
4614 ATGACGACGAGTATGATGATCCAGACAGTTGCTGAGAGGAGAAAGAGATTC 4673
4573 TCCGTCATAGGACGAGTGTATCTTCACTGTAAGAGTACACATTCATCAACGCGC 4632
4674 AGTACTCAAAAAGAAAGTGTATCTTGAAGGACCAAAATTCATCAAGACAGCA 4733
4633 AAGGACATTCGCAAAATTCATGCAATGTGCGCAAAATTCAGGCTAAATGACAGATT 4692
4734 AAGGATATGCGGAGATTAAGAGTCTGTTCCCAATGACGAGAAACGAAACGAACTG 4793
4693 TCGTTGATCTCTGGGAGAGATGATGTCAGACATCCGCTCAAAATGCGGAGAGAG 4752
4794 TGTGCTCAATATGTTGGGAGAGACCATGAAAGCAATCCGCAAAATGCGGAGAGAG 4853
4753 TCGAGGCGGTCTGCTCACTCAACCACTTCAGTCCGCTGATTAATGAGTATGACGCT 4812
4854 AACCGGTGCTAGCCCGCAAAAACGCTGCGCTGTGATGATGCTATGACATGACCA 4913
4813 GAGCGCTATATACAGTGTGCGCTGCGAAGAAAGAGATTCGCGGATGCTCATATTC 4872
4914 GAAAGGATCCACAGACTCAGAGAAATGATCAAGAAAGTAAAGTAAAGTATGCTCCAC 4973
4873 CTGTGCGGAGATCAGAGATCAGAGGCTGACAGAGATCAAGTACAGCAAAACAGTCTG 4932
4974 CCGCTTCAAGATCAAAATCAAAATGTTCAAGAGGTTCAAGTCAAAAGATGATGCTG 5033
4933 TTTTACGCGCTGATCAACGCGCTGATCAACCGGAGAGTACGCGGAA 4980
5034 TTTAACCGGATACCCCGCATTCGTTCCGCGCTGATGATCAATGAA 5081

```

```

RESULT 12
US-09-102-248-7
; Sequence 7, Application US/09102248
; Patent No. 6008035
; GENERAL INFORMATION:
; APPLICANT: Johnson, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; EXPRESSION of Heterologous Genes in the Bone Marrow
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 6008035th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,248
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/801,263
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sidley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11663 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-102-248-7

Query Match 26.7%; Score 1333.2; DB 3; Length 11663;
Best local Similarity 56.1%; Pred. No. 0;
Matches 2820; Conservative 0; Mismatches 2128; Indels 80; Gaps 13;

QY 23 ATATGAAAGAAATTCACGTTGACTTAGTGTGACAGCCGCTATGTCATGCTTACAGC 82
DB 64 AGAAGCCAGTATGATGACGTAGAGACCTCAGAGTCCGTTGTGTCATGCAACTGCAA 123
QY 83 GAGAGTTTCCACAAATTTGAGATGCAAGCAAGGAGTCACTGCAATGACCATGCCAATG 142
DB 124 AGAGCTTCCCGCAATTTGAGTGTAGTACAGAGAGTCACTCCAAATGACATGCTATG 183
QY 143 CCAAGCGTTTGGCATGTGGCAACAAAGCTCAATTGAGAGCGAAGTGCACCGGAGCAAG 202
DB 184 CCAAGCATTTTGGCATGTGGCAAGTAACTGATGCGAGTGGAGGTTCTTACCAAGCA 243
QY 203 TTATCTTGGACATTTGAAAGTGGCCCGCTGACATGACATTCATCAACCGGTATGAT 262
DB 244 CGATTTTGGACATTTGAGAGGAGGACCGGCTGTGAAATGTTTCCGAGACACAGTACAT 303
QY 263 GATATGCGCTATGATATGAGGCTGAGAGACCGGACAGACTACAAACGGTATGAGAAAGAC 322
DB 304 GCGTTTCCCGCATGCGTATGTCAGAAAGACCCGACCGCATGATGAAATATATGCGAGAAC 363
QY 323 TTAAGAAA-----GTGACATTACCGACAAAGAAATAGCTTAAAGCGGAGAGCTGC 376
DB 364 TGGCGGAAAAGCATGTAAAGATTACAAACAAAGAACTTGATGAGAAAGATCAAGGACCTCC 423

```

QY 377 TGAAGTATGTCAACACGAGAGACTCTCTGTGTATGACACAGAGCCA 436
Db 424 GGACGGTACTTGATACACCGGATGCTGAAGCGCTACTCTGTCCACACAGATGTTA 483
QY 437 CGTGTAGTACTTTGGAGTGTAGCATATACAGATGTGTAGCCAGCTGCAGCGGA 496
Db 484 CCTGACACCGCGTCCAGTACTCTGTATGACAGAGGTGTACA---TCAACGCTCCCG 540
QY 497 CATCAATACACACGAGCGCTTAAGAGTTAGGACAAATTTACATGATAGCTTTGACA 556
Db 541 GAATATTATACACAGGCTATGAAGGCGTGGACCTGTATGTAGATGTGCTTCGACA 600
QY 557 CGACCCCTTTATGTACAAAAACATGAGGTTCTTACCTTACTTACAAACAGATCGG 616
Db 601 CCACCCAGTATGTTCTCGGCTATGCGAGTTCCTTACCTGCATACAAACCACTGG 660
QY 617 CTGACGAGAGATTTGGAAGCATTAATTGCGCTCGTAACTCAGATCTTCAGAGA 676
Db 661 CCGACGAAAAAGTCTTGAAAGCGCTAACATCGACTCTGACGACAAAGCTGAGTGA 720
QY 677 GCAGCTTGAAGAAATCTCAATCTTAGAAGAAAGGCTCCAACTTACTAATAGATCA 736
Db 721 GCAGACGAGAAAGTGTGATATGAGAGAAAGAGTGAAGCCGCGGTTCACGCGTTT 780
QY 737 TATTTCTGCTGTGTTCAACATCTACAGAAAGATAGATCACTGTTAGCTGCGCATC 796
Db 781 ATTTCTCGGTGTGATGACACTTTACCGAAGACAGAGCCAGCTTGACAGACTGGCATC 840
QY 797 TTCCAAACGTGTTCACCTTGAAGAGAAAGTCTAACCTTACAGATGATGTGGACCAT 856
Db 841 TTCCATCGGTGTTCACCTTGAAGAGAAAGCTGTACCTGCGGTGTGATACAGTGG 900
QY 857 TCAGCTGTGAAGGATGATCATCAAAAGATTAAGATCAAGCCAGAGACTATACGTTAAG 916
Db 901 TGAAGTGTGAAGGCTACGTAGTGAAGAAATCAATCATGATCCCGGATTCACGGAGAAA 960
QY 917 TTGAGAACTTGGCGTCCCAATGCTCCGAGAGGTTTCTTGAAGTTGCAAAAGTCAAGATA 976
Db 961 CCGTGGGATGACCGGTTACAAACATAGCGAGGCTTTGTGTATGCAAAAGTTACCGATA 1020
QY 977 CGCTGCGGAGGAGAGGTTCTTTGTGTGTGTAGCTATGTATACAGCAACTTTGCG 1036
Db 1021 CAGTAAAGGAGAAAGGATATCGTTCCCGTGTGACGATATATCCGGCCACATATGCG 1080
QY 1037 ATCAATATACAGGATTTCTGGCACTGACGTTAGTGTGATGACCCACAAATCTATTGG 1096
Db 1081 ATCAATATACAGGATTTCTGGCACTGACGTTAGTGTGATGACCCACAAATCTTTCTGG 1140
QY 1097 TTGGGCTCAACCAAGATTTGCTGATGATGAGAGCCCAAGAAATCTAATCAATGC 1156
Db 1141 TTGGGCTCAACCAAGATTTGCTGATGATGAGAGCCCAAGAAATCTAATCAATGC 1200
QY 1157 AGAATATCTATTAACAGTGTGCGCCAGGCGTTTCCAGGTGGCGCGTGAATCTGTG 1216
Db 1201 AAAATTAATCTTCTGCAATCAATGCAAGGTTTACGAAATGGGCCCAAGAGCCCAAG 1260
QY 1217 CCGACTTGAAGAGAGAAAGATTAAGGAGTGGGAGCGCACTTTAATGAGGCTGCT 1276
Db 1261 AAGATCTTGAAGAGAAAGATTAAGGAGTGGGAGCGCAAGCTTCAATATGAGCTGCT 1320
QY 1277 GCTGGCTTGAAGAGCCAGAAATCAATCATCTAACAAGAGCTGTAGCGAAAGAA 1336
Db 1321 TGTGGGCTTTGCGACTAAGAAAGTCACTGTTTATGCGCCACCTGAAAGCGAGACA 1380
QY 1337 TTAAGAAAGTACTGCGCTCTTGTGATCTATTTGTATTCAGGCTTACAGCGACGCGG 1396
Db 1381 TCGTAAAGTCCGAGCTCTTTAGCGCTTCCCATGTCAATCCGTATGAGTACTCTTT 1440
QY 1397 TCGATATGAGGCTTCGCGGTGAGCTCAAGCTGTCTGAAACA---ACTGTCAACCCG 1453
Db 1441 TGCCCATGTCTGAGGAGAGATGAATGGAATTTGCAATCAACCAAGAAAGGAGAAAAAC 1500

QY 1454 CACCGCTATTATCAATGCGCATGTGAGCATCTGCGTGGCTTACAGCAAGAGCTGAAG 1513
Db 1501 TGCTGCAAGTCCCGAGAGAAATTAATGATGAGGCGCAAGGCTGCTTTCAGAGATGCTAG 1560
QY 1514 AAGTGTGACGCGAGAGATCAAGAGAAAGCCCTGCGCACCTTGTCTCTGAATAGAAA 1573
Db 1561 AAGATCCAGAGCGAGAGAGTCCGAGAGAGATCCCACTTATGTGACAGCAAAAGTGA 1620
QY 1574 AAGAGCCGTAGAGGAG-----AAGTGAACCTCATTTATGCAAGAGGAGAGGAG 1624
Db 1621 TCGAGGAGCTGCGAATTTGTCTGCGAATGAGAGGAGCTCCAGCGGACACCGGAGCAG 1680
QY 1625 GTAGGTGAGAGACCAAGAGAGACATCAGGTGACAAAGTTACCAGCGAGAGAGAA 1684
Db 1681 CACTGTGGAACCCCGGCGGTATGTAAAGATTAATCTTCAAGCAATATGACGATAGA 1740
QY 1685 TTGGGTCTTACGTAATCTTTCACCCAGCGGATTTGAATGAGAAAAATGCGCTGTA 1744
Db 1741 TCGGACAGTATATGCTGTCTGCGCATCTCTGTGTGAGAAAGCTTAACTCGCACAG 1800
QY 1745 TCCACCCATTGGCGGAAACAACTACTGTAATGACTCAAAAGTTAGGCGAGAGATACA 1804
Db 1801 CACACCGCTAGCAGACCGGTTTAAGATCATACGCACTCCGAAATCAGAAAGATG 1860
QY 1805 AAGTGAAGCCATPACACGCTAAGGATCATTTGACAGAGAGAGCGGCTCCCTGTTCAAG 1864
Db 1861 CAGTGTGAACATAGAGCGCTTAAATGATGATGCGACAGAGAGTCCCTTACATGCGCAG 1920
QY 1865 ACTTCCAGCATTGAATGAGAGCGCTACGATGTTTTCACAGAGAGAGGATTCGTAAACA 1924
Db 1921 AATTTAGCATGATGATGAGAGCGCCAGCTTGTGTACAAAGAAAGATTTGTAAGC 1980
QY 1925 GATACCTGACACATGSCATTCAAAGAGAGAGCGCTTAAACATGACGAAGATCTATA 1984
Db 1981 GCAAGCTGTATCATATGCTGATGACAGCGTCCCGCTAAGAAATPACAAAGAGAGCGTACA 2040
QY 1985 AGACTGTAAAGACTCAGAGACACAGTACAGATAGCTTTCGATATTGACCGACGAAGT 2044
Db 2041 AGGTTACAAAGGAGAGAGCTCCAGAGAAACAGATGATGCTTTCAGTGTGACAAAGAGCAT 2100
QY 2045 GTGTTAAGCGAAGACGAGAGTCCCTTGTGCTTAAACCGGTGATCTGATATCCACCAT 2104
Db 2101 GCGTTAAGAAAGAAAGAGCTTCAGAGACTTGTCTTTCCGAGAAAGTGCACAAACCGGCT 2160
QY 2105 TTCAAGATTTGCGTACGAGAGTCTCAAGACAGACACGACACTCAAAAGTCCCAA 2164
Db 2161 ATCAAGAACTGCTTGAAGGAGTGAAGACTCGACCGCGGTCCGTTACAAAGTTGAAA 2220
QY 2165 CCATCGAGCTATGAGAGTGCAGGTTCAAGTAAATCTGAATCATCAAAAGCGCTGTA 2224
Db 2221 CATATGAGTATAGGACACACAGATGCGCAAGTCACTATCATCAAGTCACTGTCA 2280
QY 2225 CTAAGAAAGATCTGTTGTGAGTGCAGAAAGAAATCTGCGCAAAATCATCAAGGATG 2284
Db 2281 CCGGACGATCTTGTATACAGCGGAAAGAAAGAAATCTGCGGAAATGAGCGGAGC 2340
QY 2285 TAAAGAGATAGACATGTGATGTTGCTGTAGAGCTGTGATTCAGTCTTCTTAATG 2344
Db 2341 TGCTACGCGCTGAGGAGCATGACAGATCACTGTGAAGACAGTGAATTCGTTATGCTCAAG 2400
QY 2345 GGGTTAAGCACCGGTTAACACTGTGATCTGATGAGAGCAATTTGCGGATGAGAGGA 2404
Db 2401 GATGCCAAGAGCGGTGAAGATGCTGTATGTGTGAGAAAGCGTTCCGGTCCACGAGAG 2460
QY 2405 CGCTGTGCACTGATTCATCTGTCAAAAC---TAAAGAAAGTGTATTTGTGCGGAGC 2461
Db 2461 CACTACTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
QY 2462 CAAACATGAGCGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2512
Db 2521 CTAAGCAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580
QY 2513 ATGACATATGACATGAGTGTACATTAAGATCTCTAGAGGATGACACAGACTGTAA 2572

Db 2581 AAGGATATGACCAAGACATTTCTACAAAGTTATCTCCGACGTTGACACAGCACTCA 2640
Qy 2573 CCGGCATGCTCTCCAGCCTCTTCTACGACAAAGCAATGAAACGGTTAACCCATGTGCTG 2632
Db 2641 CGGCTATGTTATCGACTGATACGATGGAAAAATGAAAAACCAAAACCCGTCAGAA 2700
Qy 2633 ATAAATCATCATGATATACCAAGGAGCCAAAGCCGCAAAAGATGATCTGATTTCA 2692
Db 2701 AGAATCATGAAATGACATTACAGGGGCCAAGAGCCGAAAGGAGGACATCATCTCA 2760
Qy 2693 CTTGTTTCAAGAGATGGGTGAAAAGGCTACAGATTGACTCAAAAATACAGAAATCATGA 2752
Db 2761 CATGTTTCCGGGTGGGTTAAGCACTGCAATCGAATCCGGACATAGAGTTATGA 2820
Qy 2753 CTGGCGTGCATCGCAAGAGATTACGGGAAAGCGTTTATGCTGTCAAGTCA 2812
Db 2821 CAGCCGGGCTCAAAAGGGCTAACAGAAAGAGATATGCGTCCGGCAAAAGTCA 2880
Qy 2813 ACGGAATCCACTCTACTCGGACATCTGAGCACTGAAAGTGTACTTACAGGACAG 2872
Db 2881 ATGAAACCCGCTACGCGATCAATCAGACATGTAACGTGTGCTCACCCGCACTG 2940
Qy 2873 AAAACGCAATGTCTGGAAGACGCTAGCTGTGATCCCTGGATAAAGCACTTACAGCTA 2932
Db 2941 AGGACAGGCTAGTATGABAACTTTACAGGGCGAACCATGATTAAGACGCTCATACG 3000
Qy 2933 AATATCCCGGGATTTTACGCGCTTCAATGAGACGTGGCAGCGGAAACGACCCATTA 2992
Db 3001 TACCTAAAGGAAATTTTCAAGGCGCACATCGAGGACTGGAAAGCTGAAACAAAGGAATTA 3060
Qy 2993 TGGCAGCGCTCTTGATTAAGCGGACAGCGTGAATGTTTCAGAAATAGGTGAACGCT 3052
Db 3061 TTGCTGCGATTAACAGTCCGCTCCGCTACCAATCCGTTAGCTGCAAGCTAACGTTT 3120
Qy 3053 GCTGGCGAAGGCTTTAGAGCCAGTCTTGGCCACGCGCCAACTGTGTGACGAGACAGC 3112
Db 3121 GCTGGGGAAGACATGGAACCGATACCTGGCCAGCGCGGTATGTACTTACCGGTTGCC 3180
Qy 3113 AGTGG--GAGAGCTTGACCCCATTTCAAGCATGACAGAGCGTACTCATTGAAATGCGAC 3169
Db 3181 AGTGAAGGACGTGTTCCACAGTTTGGGATGACAAACACACTCCGCTCATACGCT 3240
Qy 3170 TGAATCTTTTGAACCGGTTCTTTGGAGTAGACCTGGACAGGGTATTTTCCGCTC 3229
Db 3241 TAGAGTAATTTGATTAAGTTTTCGCAATGACTTACAAAGCGGCTGTTCCTCAAC 3300
Qy 3230 CTACCGTGCATTTACTTACAGGAGT-----CAGCACTGGGATA 3268
Db 3301 AGAGCATCCGTTAAGTACCATCTGCGCATCTGCGAGGCACTAGCTATTGGAGCA 3360
Qy 3269 ACTGCCAGGAGAAACATGATGGGCTTTAATAGAGATGACAAAGATTGTCAGGC 3328
Db 3361 ACAGCCCGGAAACAGCAAGTATGGGTACGATCACGCCGTTGCCCCCACTCTCCCTTA 3420
Qy 3329 GATATCCGTGATCAAAAGCGGTTGACAGAGGAGGTAGCTGATTAAGGAATATA 3388
Db 3421 GATTTCCGGTTCAGCTAGC---TGGGAAAGGACACAGCTTATTTGACAGCGGCA 3477
Qy 3389 CCATCAAGAGACTACTCTCAACAATTAATGTGTCTCAATTAATGCCGCTTGCCTCACT 3448
Db 3478 GAACATAGATTATCTCTGACACGATTAATTGTTCCAGTGAACGCAATCCTCCTCAG 3537
Qy 3449 CGTTGATGTTGACCAACAAGACAGGGTACATGATCAACGGGATTTCTATTTAAGA 3508
Db 3538 CTTTGTGTCGAGACAAAGGAAACAAACCGGCGCGATCGAAAAATTTTGAAGCCAGT 3597
Qy 3509 TGAAGGGAAATCTGTGTGTGATCGGATCTTA-----TGAGCTTCCAGG 3558
Db 3598 TCAACACCACTCCGTAATTGATCTCAGAGAAAAAATTGAAGCTCCCAAGAGAA 3657
Qy 3559 AAGAAATGATGATCGATGGGTCAATTTGCCCACTAAATCAATCAGGTGTGATCTCGATTG 3618

Db 3658 TCGAATGATCGCCCGATTTGGCATAGCCGGCCGAGATAAGAACTAACCTGGCTTTCG 3717
Qy 3619 GGAATACCTTCCATGTCGGTAAATATGACATATCTTTGTCAATGTTAGACCCGTA 3678
Db 3718 GGTTCGCGCGAGGACGG---TACAGCTGTGTATCAATATTTGAACTTAATAC 3773
Qy 3679 AGGAACCATCACTACCAACAGTCGAGATCACTATCCACCAAGCATGCTTACGTGT 3738
Db 3774 AGAACCATCACTTTCAACAGTGGGAAGACAGCGCGGACCTTGAAAACCTTTGCGGT 3833
Qy 3739 AAGCTGTCCACCACTGAACTGTGGGAAATGTGTGCTTATAGGATATGGGCTTGTCT 3798
Db 3834 TCGGCTCTGAATGCTTTAACCCCGAGGAGCCCTGTGTGAAGTCTACGGTTACGCC 3893
Qy 3799 GATTCGCCAACCGAGAAATATCATCACTGCGGTGGACGCTCATTTAGGTTACCCGCTC 3858
Db 3894 GACCGCAATGTAGAGCGATGACCGCTTGTGCAAAATTTGTCAAGATGTCTGCA 3953
Qy 3859 TGTGAGCTTAAGAACACTGCGGAAATACTGAGGTTCTTCTGTGTTCTTCCGCAAGAC 3918
Db 3954 GCGAGCCAGAGTGCCTCTCAACCAATACGAAATGTACTGATTTCCGACAACTAGAC 4013
Qy 3919 AACGCAACCAACACATGACACAGACAGACTCGGTGTAGTCTTGACAAATCTATCA 3978
Db 4014 AACGCGCACACAGCAATTCACCCGCAATCATTTGAATGTGTGATTTGCTCGGTAC 4073
Qy 3979 GGTTCACACAGTACAGAGGAGAGGCTCCAGCGTACAGAGTACAGAGTACAT 4038
Db 4074 GAGGTTCAAGAGCGAGGTTGAGCCGCGACCTGTGACCTACTTAABAGGAAACATT 4133
Qy 4039 AGCAAGAGCGCTGACCAAGCTATCTGTTAATGCTTAATAGCAAAAGTCAACAGTTCC 4098
Db 4134 GCTATTTGCAAGAGGAAGAGGATGTCAATGACGCCAATCCACTGGGACACAGAGAA 4193
Qy 4099 GGAATGTCGGTGCATCTGACCAAAATGCGCGCTCTTTGATAGACCCAAATGCT 4158
Db 4194 GGAATGTCGGTGCATCTTAATTAAGTTGCCGAAGTTTCAACCGATTACGCCACAGAG 4253
Qy 4159 GTGGGAGCGCTAGACT---TGTGAAGCAGAAACCGCTCATCATAGCTGTAGAGCC 4215
Db 4254 ACAGTACCGGAAATGACTGTGTGCGAAGAAAGAAATGATTCACGCGGTGGCCCT 4313
Qy 4216 AATTTTCTAAGATGCGGGAACCGGAGCGGACCTTAAGCTCGACGCTGCTACATGAGC 4275
Db 4314 GATTTCCGGAACACCGAGAGGACAGAGCCCTGAATTTGCTGCAAAACGCTACATGCA 4373
Qy 4276 ATACGTCATCTGTAACGCTGAGCGGATTAACAAAATATCAGTACCGCTACTCAAC 4335
Db 4374 GTGCAACATTAATTAATGAACATTAATATCAAGTCTGCGCATCCCACTGCTATTA 4433
Qy 4336 GGCATCTATTCTGTGCAAAAGATCGAGTATGCAATCATGTGCTACCTGTTCACTGCT 4395
Db 4434 GGCATTTACGAGCGGAAAGAGCGGCTTGAGGTATCACTTAATCTGTTGACAAACGCG 4453
Qy 4396 TTCACACTACGATGCGATGTCACCATATATTCTTGTGATTAACATGGGAGACAGG 4455
Db 4494 CTAGACAGACTGATGCGAGATGACATCTACCTGATGATGAAGTGAAGAGAAAGA 4553
Qy 4456 ATAAATGAGGCACTTACACCGCAAGAAAGCGT---GAATTTGATGATGACAGCA 4512
Db 4554 ATCAACGCGGTGCTCAACTTAAGAGTCTGTAACTAGCTGAAGATGAGATATGAG 4613
Qy 4513 GTAGACATGACTTGTGATGAGGTCCACCAACAGCTTTTGGCAGGACAGCGTTAC 4572
Db 4614 ATCAACGAGGATTAATGATCAATCCGACATTTGCTGAAAGGAAAGAGATTC 4673
Qy 4573 TCCGTCATGAGGCAAGTTGATTTCAATCTGTGAAGTACAGATTCATCAGCGCC 4632
Db 4674 AGTACTCAAAAGGAAGATTGTATTTGATTAAGGACCAAAATTCATCAAGAGCA 4733
Qy 4633 AAGGACATTCGGAATTCATGCAATGTGGCCCAACAATTTAGGCTTAATGACAGATT 4692
Db 4734 AAAGATATGGCGGAGATTAAGTCTGTTCCTCAATGACAGGAAAGCAAGAACTG 4793


```
1037 ATGATGACAGGATTTCTGGCACTGACGTTAGTGTGATGACGACCAAAAATCTATTGG 1096
1081 ATGATGACCGGATATATGCGCACGATATCTCACTGACGATGACCAAAAATCTTCTGG 1140
1097 TTGGGCTCAACCAAAAGATTGCTCAATGCTAGACGCAAAAGAAATCTAACCAATGC 1156
1141 TTGGGCTCAACCAAGATGCTCAATTAACGCTAGACTTAACAGAAACCAATATCATGTC 1200
1157 AGAATCTATTTACGATGCTGCGCCAGGCGCTTTCCAGGTGGGCGCGTGAATCATGTC 1216
1201 AAAATTTACTTTCGCAATCTTGCACAAAGGTTTCAGCAATTTGGGCCAAAGAGGCAAG 1260
1217 CCGACTTGACGACGAGAAAGAACTAGAGGCTGCGGAGCGCACTTAACTATGAGGCTGCT 1276
1261 AAGATCTTGACATGAAAAATGCTGGCACCAAGAGCGCAAGCTTAACTATGCTGCT 1320
1277 GCTGGGCTTTCMAAGCCGAGAAATCAATCATCTACMAAGAGCTGTGACGCAAA 1336
1321 TGTGGGCTTTCGCAATGAAAGTGCACTGTTCTATCGCCACTGGAACGCAAGCA 1380
1337 TTAAGAAAGTACCTGCGCTCTTTGACTGCTATTTGATTTCCAGCGCTTACAGCCAGGAC 1396
1381 TCGTAAAGTCCGACCTCTTTAGCGCTTTCCTCATGTCTCATCGTATGACTACTCTT 1440
1397 TCGATATGGGCTTCGCGCTAGGCTCAAGCTGCTGTTGAACCA---ACTGTCAAAACCG 1453
1441 TCGCATATGCTGCTAGGCTAGAAATGAAATTTGCAATTAACCAAGAGAGGAAAAAC 1500
1454 CACCGGCTATTAACATGCGCATGTGAGCATCTGCTGCTTACAGCAAGAGCTGAG 1513
1501 TGTGCAAGTCCCGAGGAATTAATGTAAGAGCAAGGCTGTTGAGAGATGCTGAG 1560
1514 AAGTGTGCTGACGCGGAAAGATCAAGAAAGCTTCCAGCTTGTCTCTTAATAGAA 1573
1561 AAGATCCAGAGCGGAGAGCTCCGAGAAAGCACTCCACATTAATGTGCAACAAAGTGA 1620
1574 AAGAGCGGTAGAGGAG-----AAGTAACTCATTAATGCAAGAGGCAAGAGAG 1624
1621 TCGAGGAGCTGCGAAGTTGTCTGCAAGTGAAGGAGCTCCAGCGGACACCGAGAG 1680
1625 GTACGCTGAGACACCAAGAGACATCAAGGTGACAAAGTTACCCAGCGCAAGAGAA 1684
1681 CACTGTGCAAAACCGCGCGGTATGTAAGATTAATCAAGCAAAATGACGTAAGA 1740
1685 TTGGGCTTACGCTATCTTTCAACCCAGGCGGTATTAATAGTGAATACTGCGTGA 1744
1741 TCGGACATATATGCTGTCTGCGGATCTGTGTGTAAGAAAGCTTAACTCCGACAG 1800
1745 TCGACCCATTTGCGGAAACAAATGCTGTAATGACTCAAAAGTATGAGGCGAGAGATCA 1804
1801 CACACCGCTAGCAGACCAAGTTAAGATCATTAAGCACTCCGGAAGATCAAGAAAGTATG 1860
1805 AAGTGCAGCCATACCAAGTAAAGTCAATTTGACCAAGAAAGAGCGCGCTTCCCTTTCAAG 1864
1861 CAGTGCAGACATAGAGCTAAAGTATGATGACCAAGAGAAAGTCCGTACATAGGCGAG 1920
1865 ACTTCCAGGCTTGAAGTGAAGCGGTACGATGTTTTTAACGAGAGGAGTTTCTTAACA 1924
1921 AATTTTACGACTGATGAGAGCGCGCTTGTGTACAAAGAAAGATTTTGTGAACC 1980
1925 GATGCTGACCAACATGCGATCAACGAGAGAGGCTTAAACATGACGAAAGTATCTTA 1984
1981 GCAAGCTGTACCATATTCATGACGAGTCCCGCTTAAGAAATACAAAGAGAGGAGTACA 2040
1985 AGACTGTAAAGACTCAGACACAGATCAGAAATAGTCTTGATTTGATACGCAAGAAAGT 2044
2041 AGGTATCAAAAGGAGAGCTCCGAGAAACAGATAGTGTGACCTGACAAAGAGGAT 2100
2045 GTGTAAAGCGAAGAGCGAGGTCTTGTGCTTAACCGGTGATGTGTAATCCACAT 2104
2101 GCGTAAAGAGAGAGAGCTTCAGAGATTTGCTTTCGGGAGAACTGACCAACCGGCT 2160
2105 TTCAAGATTTTGCAGAGAGTCTCAAGACACGACGACGACCTCAAAAGTCCAA 2164

2161 ATACGAACTAGCTCTTGAGAGACTGAAAGCTGACCGCGGTCCGCTACAAAGTTGAAA 2220
2165 CCAATGAGTCTATGAGTCCAGGCTTCAAGTAAATGGAATCACTAAAGAGCTTGA 2224
2221 CAATGAGAGTATGAGACACAGAGATCGGCAAGTCAAGTATCTATCACTCACTGCA 2280
2225 CTAAAGAAAGATCTGTTGTGATGTCGAAAGAGAAAGAACTCGGCAAAATCATCAGGATG 2284
2281 CCGCAGGTGATCTTGTATTCAGCGGAAAGAAAGAAATCTCCGGAATTTGAGGCCGACG 2340
2285 TAAAGAGATGAGACGATGATGTGCTGCTAGAGCTGTGATTCAGTCTTAAATG 2344
2241 TGTAGCGGTGAGGAGGATCAGATCAGTCAAGAGAGTGAATTCGTTATGCTCAAG 2400
2345 GGGTTAAGCACCCGTTAAACATCTGACATTTGATGAGGATTTGCTGCAATCAGAGGA 2404
2401 GATGCCCAAAAGCGGTAGAGATGCTGATTTGACGAAAGGCTTCGCGTGCACAGAGAG 2460
2405 CGCTGCTGGACATGATGCGCATGTCAAAC---TAAAGAAAGTATGTTGCGGAGAC 2461
2461 CACTACTTGCTTGTATTCATATGTCAGACCCGCTAAGAAAGTATCTATGCGGAGAC 2520
2462 CAAACCAATGCGGCTTCTTTAAGATGATGCTGAAAGTACATTTTACC----- 2512
2521 CTAAAGCATGCGATCTTCAAGATGACATTAAGTAAATTTCAACACACCTGAAA 2580
2513 ATGACATATGACTGAAGTGTACATTAAGATCTTTAGAGGTGACACACATGTA 2572
2581 AAGCATATGTACCAAAACATTTCAAGTTTATCTCCAGCTTGCACACAGCACTCA 2640
2573 CCGCATGCTCTCAGCTCTTCTAGACAAAGGATGAAGAGGTTTAAACCATGCTG 2632
2641 CCGCTATTTGATGACACTGCTATTCATGAAATGAAATCCAAACCCGTGCAAG 2700
2633 ATAAATCATATGATATACCAAGGACCAAAAGCCGCAAAAGATGATCTGATTTAA 2692
2701 AGAATCATGAATGACATTAACAGGCGCAAGAGCGGAGCATCATCTCTGA 2760
2693 CTTGTTCAAGAGATGGGTGAACAGTACATGATTTGCTCAAAATCAAGAAATCTATGA 2752
2761 CATGTTTCCGCGGTGTTAGCAATGCAATCCGACATGAGGATTAATGA 2820
2753 CTGCGGCTGATGCAAGACTTACGCGGAAAGGCGTTTATGCTGCTAGGTAACAAATGA 2812
2821 CAGCGCGGCTTCAAGAGGCTAACGAAAGAGATATGCTGCTCCGCAAAAGTCA 2880
2813 ACGAAATCACTACTCTGACAGCTTCTAGACAGTGAACGTTTCTTACAGCAG 2872
2881 ATGAAACCGGCTGTACGCAATCATGAGGATGTAAGTGTGCTCAACCGCACTG 2940
2873 AAAAAGCATTTGTGGAAGACGCTAGCTGTATCCCTGATTAAGACACTTACACTA 2932
2941 AGGACAGGCTAGTATGAAAACTTTACAGGCGGACCATGATTAACAGCTCTCAAG 3000
2933 AATATCCCGGGGATTTACAGGCTTCAATGACGCTGCGAGCGGAGACAGACGCAATTA 2992
3001 TACTTAAAGGAATTTACAGGACCATCGAGACTGGAGAGTGAACAGGAAATTA 3060
2993 TGGCAGCGTCTTGTATGAGCGGACAGCTGATGTTTCAAGATTAAGTGAAGCTCT 3052
3061 TTGCTGGATTAACAGTCCGCTCCCGTACAAATCGCTTCAAGTGAAGCTAACGTTT 3120
3053 GCTGGGCGAAGCTTTAGAGCACTTGTGCGCAGCGCAACATTTGCTGACGAGACAC 3112
3121 GCTGGGCGAAGCACTGGAACCGATACCTGCGCAGCGCGGATGCTTACCGGTTGCC 3180
3113 AGTGG---GAGACCTTGAACCATTAAGCATGACAGAGGTACTCACTGAATATGCGAC 3169
3181 AGTGAACGAGCTGTCCACAGTTTGCAGTGAACCAACATCTGCGCATCTACGCT 3240
3170 TGAACCTCTTTGACACAGGTTCTTGTAGATGACAGTGAAGTATTTTCCGCTC 3229
```

Db 3241 TAGAGTAATTGCAATTAAGTTTTCGGAGTAGACTTGAACAGCGGCTGTTTTCAAAC 3300
Qy 3230 CTACCGTCGCACTTACTTACAGGAT-----CAGCACTGGGATA 3268
Db 3301 AGAGCATCCGTTAAAGTACATCTCTGCCACTACGCGAGCCAGTAGCTCACTTGGAGCA 3360
Qy 3269 ACTCCGCAAGGAAGAACTGTATGGGCTTAAATAGAGAGGTAGCAAAAGAGTGTACGGC 3328
Db 3361 ACAGCCGAGGAACAGCAAGTATGGGTACGATCAGCGGTTGCCCGCAACTCTCCGTA 3420
Qy 3329 GATATCCGTGATCAACAAAAGCGTTGACACAGCGAGGTAGCTGATATAGAAATAA 3388
Db 3421 GATTTCCGGTGTTCAGCTAGC---TGGGAAAGGACACAGCTTATTTGACAGGGGA 3477
Qy 3389 CCATCAAGGACTACTCTCCAACTAATATGTGTTCCATTAAATGGCGGTTGCCCACT 3448
Db 3478 GAACATAGATTAATCTCTGACAGCATTAATTTGTCTCCAGTAAACGCAATCTCTCCAG 3537
Qy 3449 CGTTGATCGTTGACCAAAAGGACAGGTAACATGATCAACGCGATTTCTATCTAAGA 3508
Db 3538 CTTATGTCGGGACAGCAAGGAAACAACCGGCGGTCGAAAATTTCTTGAGCCAGT 3597
Qy 3509 TGAAGGCAATCTGTGTGTTGATCGGCGATCTA-----TCAGATTCAGAGG 3558
Db 3598 TCAACACACACTCCGTACTTGTGATCTCAGAGAAAAAATTGMACTCCCAAGAGAA 3657
Qy 3559 AAGAAAGTAGTCATGGGTCAATTTGCCCACTAATACATAGGTGATCTCATTTG 3618
Db 3658 TCGAATGATGCCCCCGATTGGCATAGCCGGGACAGATAGACATCAACCTGGCTTCG 3717
Qy 3619 GGAATACCTAGCATATGCGTAAATATGACATTAATTTGCAATGTTAGACCCCGTAC 3678
Db 3718 GGTTCGCGCGAGGACG---TACGACTGGTTCATCAATATTTGAACTAAATAC 3773
Qy 3679 AGAAACATCACTACCAACAGTGCAGAGATCAGCTATTCACACAGATGTAACTGT 3738
Db 3774 AGAAACATCACTTCAACAGTGCAGAGACACGCGGACCTTGAACCTTTGGGCT 3833
Qy 3739 AAGGTGTCACCACTGAACACTGGCGGAACATGTGGCTATAGGATAGGGCTTCT 3798
Db 3834 TGGGCTGTAATGCTTTAAACCCGAGGACCTCTGAGTGAAGTCTTACGTTACCC 3893
Qy 3799 GATCGCGCAACGAGATATCATCTGCGGTGACCTCATTTAGGTTTACCCTGTC 3858
Db 3894 GACCGCATATGAGAGAGTACTACCGCTCTTGCAGAAAAATTTGACAGATGTCGCA 3953
Qy 3859 TGTCAAGCTTAAGAACACTGCGCAAAATCTAGAGTCTCTTGTGTTCTTGGCAGAGAC 3918
Db 3954 GCGAGGCCAGAGTGGCTCTCAAGCAATACAGAAATGTACTGATTTTCCGACACTAGAC 4013
Qy 3919 AACGGCAACACACATGACCAAGACAGATCGGTGATGCTTGACAACTATACAA 3978
Db 4014 AACAGCCGACACGCAATTCACCCCGATCATTTGAATGTGTATTTGCTCCGTGAC 4073
Qy 3979 GGGTCAACCAAGTACAGGACGAGAGAGTCCAGGCTACAGAGTATCAGAGGTACAT 4038
Db 4074 GAGGTTAAGAGAGAGAGTGTGAGCCGACCTGTGTACCTTAATTAAGAGAACTTT 4133
Qy 4039 AGCAAGAGCGCTGACCAAGTATCTGTAATGTGTGTAATAGCAAAAGTCAACAGGTTCC 4098
Db 4134 GCTGATTTCAAGAGAGAGAGTGTGTCAATGACCAATTCACCTGGGACACAGAGAA 4193
Qy 4099 GGAAGTCCGCTGACCTGTACCGAAAAATGCGGCTGCTTTGATAGACAGCAATACT 4158
Db 4194 GGAAGTCCGCTGATTAATAAGTTGGCGCAACAGTTTCAACGATTCAGCCACAG 4253
Qy 4159 GTCCGAGCGCTAGACT---TGTGAAGCAGAAACGCTCATCATCATGCTAGAGACC 4215
Db 4254 ACAGGTACCGCAAACTGACTGTGTGCAAGAGAAAGATGATCAAGCGGTGGCCCT 4313
Qy 4216 AATTTTCTAAGATCCGGAACCGGAGGCGCACTTAAAGCTGACGCTCTACATGAGC 4275
Db 4314 GATTTCCGGAACAACCCAGAGGCAAGAGCCCTGAATTTGCTGCAAAAGCCCTACATGCA 4373

Qy 4276 ATAGGTCATGTCACAGCTGAGGAGTTCAAAAATATGAGTACCGCTACTGTCAACC 4335
Db 4374 GTGGCAGCTTAATTAATTAATATATCAAGTGTGCGCATCCACTGCTATATCA 4433
Qy 4336 GGCATTAATTTGTGTCGCAAGATGAGTATGATGATCATTCATCACTGTTCACTGCT 4395
Db 4434 GGCATTTACGAGCGGAGAAAGACCGCTTAGGATACATTAATCTGTTACAAACCGG 4493
Qy 4396 TTGCACATTACGATGCGGATGTCACCTATATTTGTTGATTAACAATGGAGACCG 4455
Db 4494 CTAGACAACTGATGCGGAGTAACTTACTGCTCGATTAAGAGAGTGAAGAGAA 4553
Qy 4456 AATAAGAGGCAATTCACCGCAAGAAAGCGTC---GAAATTCGATGATGACAAAGCA 4512
Db 4554 ATGACGCGGTGCTTCACATTAAGAGTCTTAATGAGCTGAAAGATGAGATATGAG 4613
Qy 4513 GTAGACATTGACTTGTGTCAGGTCACCCAAACAGCTTTTGGCAGGACGACAGGTTAC 4572
Db 4614 ATGACGACGAGTATGATGATCCATCCGACAGTTGCTGAAAGGAAAGAGGATTC 4673
Qy 4573 TCCGTCATGAGGCAAGTGTATTCATACCTGGAAGTACAGATTCATCAACCGCC 4632
Db 4674 AGTACTCAAAAGAAAGTGTATTCGTACTTGAAGGACCAAAATTCATCAAGCAGCA 4733
Qy 4653 AAGCATTTGCCGAATCCATGCAATGCGCCCAAAATGAGGCTAATGAGCAGATT 4692
Db 4734 AAGATATGCGGAGATTAAGTCTCTGTTCCAAATGCAAGAAAGCAAGCAACATCG 4793
Qy 4693 TGCTTATCATCTGCGGAGATATGTCAGCATCCGCTCCAAATGCGCCAGTAGAGAG 4752
Db 4794 TGTGCTATCATATTTGGGAGAACCATGGAACATCCGCGAAATGCGCGGTGACAC 4853
Qy 4753 TCAAGGCGTGTCTCCACTCAACACTTCCATGCTGTGTAAATACCTATGACGCT 4812
Db 4854 AACCGTGTCTACCCGCAAAACGCTGCGTCTGTATGATGATGATGATGATGATG 4913
Qy 4813 GAGGCGCATACAGTGGCTGCGGAGAAAGAAAGTTCGCGTATGCTCATATTC 4872
Db 4914 GAAAGGTCACACATCAGAAAGATTAACCTCAAGAAATTAAGTATGCTCTCCAC 4973
Qy 4873 CTGTTGCCAGTACAGATCAGGCGTGCAGAGATCAGTACAGCAACAGTCTCG 4932
Db 4974 CCCCTTCAAGTACAAATTAAGATGTTAAGAAAGTTCAGAGTTCACAAAGTACTG 5033
Qy 4933 TTTTCAAGCGTGTACCAACGCGCTGTACACCCAGAAATGACCGGAA 4980
Db 5034 TTTAACCGCATACCCCGCATTCGTTCCCGCTGAATGATACATAGAA 5081

RESULT 14
US-09-367-764-7
Sequence 7, Application US/09367764
Patent No. 6583121
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Selitzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 583121th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/367,764
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEO ID NO: 7
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-367-764-7

Query Match 26.7%; Score 133.2; DB 4; Length 11663;
Best Local Similarity 56.1%; Pred. No. 0;
Matches 2820; Conservative 0; Mismatches 2128; Indels 80; Gaps 13;

QY 23 ATATGAAAGATTCACTTACCTTACCTGACAGCCGCTATGCTCAAGCTTTACACG 82
DB AGAAGCCAGTGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 123
QY 83 GAGAGCTTCCAGATTTGAGATCGAAGCAGAGCAGAGTCACTGACATGACATGACATG 142
DB AGAGCTTCCGCAATTTGAGATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 183
QY 124 AGAGCTTCCGCAATTTGAGATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 183
DB 143 CCAGAGCGTTTCCAGATTTGAGATCGAAGCAGAGTCACTGACATGACATGACATGAC 202
184 CCAAGAGCTTTCCAGATTTGAGATCGAAGCAGAGTCACTGACATGACATGACATGACATG 243
QY 203 TTATCTTGAATGGAAGTGGCCGCTGACATGACATGACATGACATGACATGACATGAC 262
DB 244 CGATTTTGAATGGAAGTGGCCGCTGACATGACATGACATGACATGACATGACATGACAT 303
QY 263 GTATCTGCTTATGAAAGTGGCCGCTGACATGACATGACATGACATGACATGACATGACAT 322
DB 304 GCGTTTGGCCGCTTATGAAAGTGGCCGCTGACATGACATGACATGACATGACATGACATG 363
QY 323 TTAAGAAA-----GTGACATTTACCGAAGAACTAGCTCTTAAGGCGGACAGCTGC 376
DB 364 TGGGGAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 423
QY 377 TGAAGTCAATGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 436
DB 424 GGACCGTAACTTGAATGCAACGACGACGACGACGACGACGACGACGACGACGACGACGAC 483
QY 437 CGTGAAGTCACTTTGAAAGTGGAGATGATCAAGATGATGATGATGATGATGATGATGATG 496
DB 484 CTTGCAACACGCGTCCGAGTACTGCTGACGACGACGACGACGACGACGACGACGACGACG 540
QY 497 CATCAATCTACACGAGCGCTTAAAGAGTTAGGACATTTACTGATAGCTTTTGAACA 556
DB 541 GAACTATTTACACAGGCTATGAAAGGCGTGGGACCTGATGATGATGATGATGATGATGATG 600
QY 557 CGAGCCCTTTATGATGACAAAAACATGAGGTTCTTACCTTACTTACCAACGAACTGGG 616
DB 601 CCACCCAGTTCATGTTCTCGGCTATGAGGATGCTTACCTGACATCAACCACTGAGG 660
QY 617 CTGACGAGAGGATTTGAGAGCAGTAACTTGGCTGATGATGATGATGATGATGATGATGATG 676
DB 661 CCGAGAAAAGTCTTGAAGGCGTAACTGAGCTTGGACGACAAAGCTGAGTGAAG 720
QY 677 GCAGGCTTGAAGAACTCTCAATCTTTAGAGAAAGAGGCTCAACTACTAATAGATCA 736
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 721 GCAGACGAGAAAGTTGCTCCATATGAGAAAGAGGTTGAAGCCCGGCTCACGGCTTT 780
QY 737 TATTTCTGTTGTTGTTCAACAATCTTACACAGAGATGATCACTGTTACGTTAGTGGCATC 796
DB 781 ATTTCTCGTTGATGACACTTTTACCCAGAACACAGAGCAGCTTGCAGAGCTGGGATC 840
QY 797 TTCCAAACGTGTTTCACTTGAAGAAAGTCTTACTTCAACAGGTAGATGTTGGACCAATG 856
DB 841 TTCCATCGGTGTTTCACTTGAAGAAAGCAGTGTACTTCCGCTGTGATGATGAGTGG 900
QY 857 TCAGCTGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 916
DB 901 TGAGCTGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 917 TTGAGACTTGGGGTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 976
DB 961 CCGTGGATACGGGTTTCAACAATGACAGGCTTCTTGTATGATGATGATGATGATGATGATG 1020
QY 977 CGCTGCGGCGAGAGGCTTTCTTGTGCTGTGTGATGATGATGATGATGATGATGATGATGATG 1036
DB 1021 CAGTAAAGAGAAACGGGTATCGTCCCGGTGACATGATGATGATGATGATGATGATGATGATG 1080
QY 1037 ATCAGATGACAGGATTTGAGCACTGACCTTACTGATGATGATGATGATGATGATGATGATG 1096
DB 1081 ATCAGATGACAGGATTTGAGCACTGACCTGATGATGATGATGATGATGATGATGATGATG 1140
QY 1097 TTGGGCTCAACCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1156
DB 1141 TTGGGCTCAACCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
QY 1157 AGAATATCTTATTAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1216
DB 1201 AAAATATCTTATTAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY 1217 CCGACTTGAACAGAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1276
DB 1261 AAGATCTTGAACAGAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1320
QY 1277 GCTGGGCTTCAAGACCCAGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1336
DB 1321 TGTTGGGCTTCCGCAATTAAGATGACATGCTTCTATGCTCCACATGAGAAACGAGACCA 1380
QY 1337 TTAAGAAAGTACCTGCTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1396
DB 1381 TCGTAAAGTCCAGGCTCTTTTAAAGCTTCCCATGATGATGATGATGATGATGATGATGATG 1440
QY 1397 TCGATATGAGGCTTCCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1453
DB 1441 TGCCATGCTGCTGATGAGCAGAAATGAAATGAGCATTACCAACCAAGAGAGAGAAAC 1500
QY 1454 CACCGGCTATTAACATGAGCCGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1513
DB 1501 TGCTGCAAGTCCGAGAGAAATGATGATGAGAGCCAGAGCTGCTTGTGAGAGATGCTCAG 1560
QY 1514 AAGTGGCTGACGCGAAGATGACAGAAAGCCCTGACCTTGTCTCTGAAATAGAAA 1573
DB 1561 AGAATCCAGAGGAGAGAAAGTCCGAGAACATCCACATTTAGTGGACAGACAAAGTGA 1620
QY 1574 AAGAGCCGTGAGAGCAG-----AAGTAACTCTATTATGCAAGAGGACGAGACAG 1624
DB 1621 TCGAGAGAGCTGCGAGAGTGTCTGCGAAGTGGAGGCTCCAGGCGAGACACCGAGACG 1680
QY 1625 GTACGCTGAGAGACCAACGAGGACATCAAGGGTGAACAAGTTTCCAGGCGAAGAGAGA 1684
DB 1681 CACTGTGAGAAACCCCGCGGTGATGATGAGATTAATCTCAAGCAATGACCGTATGA 1740
QY 1685 TTGGGCTTATGCTATCTTTCACCCAGGCGGATTTAGTAAAGAACTGGCGTGA 1744
DB 1741 TCGGACAGTATATGTTGTTCTCGCCGATCTGTGCTGAGAGAGACGTAATCTGCACAG 1800
QY 1745 TCACCCATTTGGCGGAACAAGTAACTGATGATGATGATGATGATGATGATGATGATGATGATG 1804
DB 1801 CACACCGGTAGCAGACCAAGTTAAGATCAATACGCACTCCGGAAGATCAAGAAAGTATG 1860
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

1805 AAGTCAGCCATACACGCTAAGTCAATTGTACAGAAAGGACGGCGCTCCCTGTTCAAG 1864
1861 CAGTCGAAACCATACGACGCTTAAGTACTGATCCAGCAGGAAGTCCCGTACCATGCGCAG 1920
1865 ACTTCAGGCAATTGAGTGAAGCGCTAGCATCGTTTCAAGAGAGAGGTTTGTAAACA 1924
1921 AATTCTTAGCAGTGTGAGAGCGCCACGCTTGTGTACAAAGAAAGAGTTTGTGAACC 1980
1925 GATTACCTGACCAATCCGATTCACAGAGAGCGCTTAACAATCAAGAGAGTACTATA 1984
1991 GAAAGCTGACCAATATGCTACAGACGCTCCGCTAAGATACAGAAAGAGAGCATTA 2040
1985 AGACTGTAAAGACTCAGACAGACAGACTCAGAAATAGCTTTCATATTTGACGACGAAAGT 2044
2041 AGGTTACAAAGGCAAGAGCTCGAGAAACAGAGTACTGTGTAAGTGTGACAAAGAGCAT 2100
2045 GTGTTAAGGAGAGACGCGAGGTCCTTGTGCTTAACCGGTATCTGTAGATCCACAT 2104
2101 GCGTTAAGAGAGAGAGGCTCAGAGCTTGTCTTTCGAGAGATGACCAACCGCCCT 2160
2105 TTCAAGAGTTTCCGATACAGAGTCTCAGACAGACAGACAGACTTCAAGAGTCCCA 2164
2161 ATCAGAACTAGCTCTTGAAGGATGAAAGCTGACCGCGGTCCCGTACAAAGTTGAAA 2220
2165 CCATCGAGTCTATGAGTGCAGGTTCCAGTAAATCTGGAATCAATCAAAAGCGCTGTA 2224
2221 CAATAGAGATGATAGGACACACAGAGTCCGCAAGTCACTATCATAGTCAACTGTCA 2280
2225 CTAAAGAAAGATCTGTTGTGAGTGCAGAGAGAAAGAACTGCGCAGAAATCATCAGGATG 2284
2281 CGGCACTGATCTTGTACCAAGGAAAGAAAGAAAGAAAGTGAAGGCGGAG 2340
2285 TAAGAGAGATGAGATGATGATGTTGCTGCTAGAGACTGTCATTCAGTCTTAAATG 2344
2341 TGTACGCGTGAAGGAGCATGAGATCACTCGAAGACAGTGAATTCGTTAGTCAACG 2400
2345 GGGTTAAGCAACCCGCTTAAACCTGTGATGATGAGGCAATTCGCTGCCATGACGAGA 2404
2401 GATGCCAAGAGCCGTAGAGAGCTGTATGTTGAGAGAGCGTTCCGGTCCACGAGAG 2460
2405 CGCTGTCGACATGATGTCATGCTCAACCC--TAAGAAAGTGTATTTGTCGGGAGC 2461
2461 CACTACTGCTTGTATGCAATGTCAGACCCCGTAAAGAGTATGATTCGGAGACC 2520
2462 CAAAACATGCGGCTTTTAAATGATGATGCTGAAAGTATTTTAAAC----- 2512
2521 CTAAACATGCGGATTTCTTCAACATGATGCACTAAAGGTACATTTCAACCCCTGAAA 2580
2513 ATGACATATGACCTGAATGATCACTTAAGAGATCTAGAGGTGCAACAGACTGTAA 2572
2581 AAGACATATGATCAAGACATTTCTAAGTTTATCTCCGAGCTTGCAACAGCCAGTCA 2640
2573 CCGCATGCTCTCAACGCTCTTCTACGACAAAGAGATGAGAGCGGTTAACCATGTGCTG 2632
2641 GCGCATATGATCGACATGATGATGAGTGAAGAAATGAAACCAAAACCCGTGCAAGA 2700
2633 ATAAATCATCATAGATACCAAGAGACCAAAAGCCGCAAAAGATGATCTGATTTAA 2692
2701 AAGACATGAAATCAATTAAGAGGAGCCAGAACCCGAGCAGATCATCTGTA 2760
2693 CCGTTTGAAGAGATGGGTAAGCAGCTACAGATTGACATCAAAATATCAGAAATCAAGA 2752
2761 CATGTTTCGCGGGGTTAAGCACTGCAATGATGATCTCCGAGCATGAGGTAAGA 2820
2753 CTGCGGCTGATCGCAAGACTTACCGGAAAGGCGTTTATGCTGTCAAGTCAAGTCA 2812
2821 CAGCGCGGCTTCAAGAGGCTTACAAAGAAAGGATATGCTCGTCCGGCAAAAGTCA 2880
2813 ACGAGATCTACTTACTGCAAGCTTTGAGACAGTGAAGCTGTTACTTACAGCAAG 2872
2881 ATGAAACCCGCTGTACGAGATCAGATGAGACATGTGAAGCTGTTGCTCACCCGACATG 2940

2873 AAAAAAGCATTTCTGGAAGACGCTAGCTGTGATCCCTGATTAAGACATTAAGCTTA 2932
2941 AAGACAGGCTGTATGAAAACTTTACAGGCGAGACCATGATTAAGCAGCTCAATACG 3000
2933 AATATCCGGGAGATTTACGCGCTTCATTTGACAGACTGACAGCGGCAACAGCCATTA 2992
3001 TACCTAAAGAAATTTTCAGGACCAATCGAGAGCTGGGAAGCTGAACAAAGGAATTA 3060
2993 TGGCAGCGGTTCTGTATAGCGCGACAGAGCTGATGATGTTTCCAAATTAAGTGAAGCT 3052
3061 TTGCTGCAATTAACAGTCCCGCTCCCGTACCAATCCGTTACGCTGCAAGACTTAAGTT 3120
3053 GCTGGCGAAGGCTTTAAGCAGATCTTGCGCACCGGCCAACTTGTGCTGACAGACAGC 3112
3121 GCTGGCGAAGGCACTGGAACCGAATCTGCGCACCGCGGATGCTTACTTACCGTTGCC 3180
3113 AGTGG--GAGAGCTTGACCCATTCAGATGACAGAGCGTACTACCTGAATGCGAC 3169
3181 AGTGAAGGAGCTGTTCCACAGATTGCGGATGACAAACACACTCGGCCATCTACGCT 3240
3170 TGAATCTTTTGGACAGGTTCTTTGAGATGAGACTGGAAGATGATTTTCCGCTC 3229
3241 TAGAGTAATTGCAATTAAGTTTTCGCGATGAGACTTGAACAAAGCGGCTGTTTTCAAAC 3300
3230 CTACCGTGCATTAATTACAGAGAT-----CAGCACTGGGATA 3268
3301 AGAGATCCCGTTAAGTATCATCTGCGCATACGACAGCCAGTACTCATTTGGACA 3360
3269 ACTGCGAGGAGAGAACTGTATGCGCTTAAATAGAGATGACAAAGAGTTGTCAACGCG 3328
3361 ACAGCCAGAGAAACAGCAATGATGGGTACGATCAACGCGTTGCGCGCAACTCTCCGTA 3420
3329 GATATCCGTCATCAAAACCGTTGACAGAGAGGATAGCTGATTAAGAAATATA 3388
3421 GATTTCCGCTGTTCCAGCTAC--TGGAAGAGGACACAGCTGATTTGACAGAGGCA 3477
3389 CCATCAAGAGCTACTCTCCAAATTAATGATGTTTCAATTAATGCGCGTTGCCCACT 3448
3478 GATCTAGATATCTCTGACAGCTTAATCTTGTCCTGTAACCGCAATCTCTCCACG 3537
3449 CGTTGATGTTGACCAAAAGAGCAGGTTACATGATCAAGCGGATTCCTATCTAAGA 3508
3538 CTTATGTCGCCGAGACAAAGAGAAACAAACCGCGCCGCGTGAATAATCTTTGAGCGAT 3597
3509 TGAAGCGAAATCTGTGTGATGATCGGATCTTA-----TAGCATTCAGAG 3558
3598 TCAACACACCTCCGTAATCTTGTGATCTCAGAGAAATAATGAAGCTCCCAACAGAGA 3657
3559 AAGAAAGTATGATCATGCGTCCATGTCACCTAATACATCAGGTGATCTGATTTG 3618
3658 TCGAATGATGCGCCGATTTGCGCATAGCGCGGAGATTAAGAACTAACCTGCTTTCG 3717
3619 GAAATACCTTACCATGTCGTAATATGACATTAATCTTTGCAATGTTAGACCCCGTAC 3678
3718 GGTTCGCGCGAGGACG---TAGACCTGAGTTTCAATCAATATGGAATGAATTAATAC 3773
3679 AAGAACTATCTACCAACAGTGCAGAGATACGCTATCAACACAGATCTGAACGTGT 3738
3774 AAGAACTATCTTCAACAGTGCAGAGACCAAGCGGAGCACTTGAATAACCTTTGCGGT 3833
3739 AAGGCTGTCACACCTGGAACACTGCGGAGAACATGTGAGGCTATAGGCTATAGGCTTCT 3798
3834 TCGGCTCTGAATGCTTAAACCGAGAGAGCCCTCGTGTGAGAGTCTTACGCTTACGCC 3893
3799 GATGCGCAACCGAATATCATCTGCGGTGACGCTCATTTAGGTTTAAACCGGTTC 3858
3894 GACCGCAATATGAGAGATGATGATCAACGCTCTTGCAAGAAATTTGTAGAGTGTGCA 3953
3859 TGTACCTTAAGAACTGCGGAAATATAGAGGTTCTCTTGTGTTTTCGCGAGAGAC 3918
3954 GCGAGGCGAGAGTCTGCTCAAGCAATACAAATATGATCTGATTTTCCGAACTACAC 4013
3919 AAGCGCAACCAACATGACAGAGACAGACTCGGTGATGCTTGAACATCATATCA 3978

```

Db      4014 AACAGCGGACACGACCAATTCACCCCGCATTTGATGATTTGTCCTGCTAC 4073
Qy      3979 GGGTCACACAGSTACGAGCGGAGAGCTCCAGCTTCAAGATGATCAGAGTGACATT 4038
Db      4074 GAGGATCAAAAGACGGAGTTGGAGCCGACCGTGATCCCTACTTAAAGGAGAAACATT 4133
Qy      4039 AGCAAGAGCGGTGACCAAGCTATGCTTATGCTGTAATAGCAAGGTCACCAAGTTCC 4098
Db      4134 GCTATTTGTCAAGAGGAGAGAGCTGTCAATGACCCATTCACCTGGGAGCCAGAGAA 4193
Qy      4099 GGAGTGTGCGGTGCACTGTACCGAAATGCGCGCTGCTTTGATGACACCAATAGCT 4158
Db      4194 GGAGCTGCGCGTGCATCTATTAAGTTGGCCGAACAGTTTCACCGATTGACCAAGAG 4253
Qy      4159 GTGGGAGCGGTAACT---TGTGAAGACGAAACCGCTCATCATCATGCTGTAGACCC 4215
Db      4254 ACAGGTACCGGAAATGACGTGTGCGCAAGAAAGAAAGATCCACGCGTGGCCCT 4313
Qy      4216 AATTTTCTAAGATGCGGAAACCGAGGCGCACTTAAGTGGAGCTGCTACATGAGC 4275
Db      4314 GATTTCGGGAAACACCCAGAGGACAGAACCTTGAAATTTGCTGCAAAAGCCTTACATGCA 4373
Qy      4276 ATAGCGTTCATGTCACACGCTAGCGGATTAACAAATATGATGACCGTACTGTCAAC 4335
Db      4374 GTGCGAGACTTAAATGAACATTAATCAAGTCTGTGCGCATCCCATCTATCTACA 4433
Qy      4336 GGCATCTATTTGCTGGCGAAAGATGAGTATGATGATGATGATGATGATGATGATGATG 4395
Db      4434 GGCATTTAAGCGAGCGGAAAGACCGCTTGAAGTATCTACTTAATGCTTGAACACCGCG 4493
Qy      4396 TTCACACTAGAGTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 4455
Db      4494 CTAGACAGAACTGATGCGGAGTGAACATCTAATCTGCTGATGATGATGATGATGATGATG 4553
Qy      4456 ATAAATGAGGCAATTCACCGGAAAGAAAGCGTC---GAAATTCGATGATGATGATGATG 4512
Db      4554 ATGACGCGGCTGCTCAACTTAAGAGTCTGTAATCTGATGATGATGATGATGATGATG 4613
Qy      4513 GTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4572
Db      4614 ATGACGAGCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4673
Qy      4573 TCCGTCATGAGGCGCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4632
Db      4674 AGTACTACAAAAGAAAGTTTATTTCTGATGATGATGATGATGATGATGATGATGATGATG 4733
Qy      4633 AAGGACATTTGCGGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 4692
Db      4734 AAGATATGCGGAGATTAAGTCTGTTCCCAATGACAGAAAGCAAGCAACTG 4793
Qy      4693 TGCCTTATCATCTGCTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4752
Db      4794 TGTGCTTACATATTTGGGGAGACCATGAGCAATTCGCGGAAAGATGCGCGTGGACAC 4853
Qy      4753 TCAGAGGCGTGTGCTTCACTCACACCTTCATGATGATGATGATGATGATGATGATGATG 4812
Db      4854 AACCGTGTGTTAGCCCGCAAAACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4913
Qy      4813 GAGCGCTATTAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4872
Db      4914 GAAAGGGTCCAGACCTCAGAAAGCAATTAAGTCAAAAGAAAGTAAAGTAAAGTAAAGTAA 4973
Qy      4873 CTGTGTCGGAAGTACAGATACAGGCGTGCAGAAAGTACAGTACAGAAACAGTCTG 4932
Db      4974 CCCCTTCCAAAGTCAAAATTAAGATGTTCAAGAAAGTTCAGTCACAAAAGTATGCTCG 5033
Qy      4933 TTTTACAGCGGTGATACACCGGCTGTACACCGAGAAAGTACGCGGAA 4980
Db      5034 TTTTAAACCGCATACCCCGCATGCTGTTCCGCGCTGATGATGATGAA 5081

```

```

US-08-801-263A-8
: Sequence 8, Application US/08801263A
: Patent No. 5811407
:
: GENERAL INFORMATION:
: APPLICANT: Johnston, Robert E.
: APPLICANT: Davis, Nancy L.
: APPLICANT: Simpson, Dennis A.
: TITLE OF INVENTION: System for the In Vivo Delivery and
: NUMBER OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
: CORRESPONDENCE ADDRESS:
: ADDRESS: Bell Seltzer Park & Gibson, P.A.
: STREET: 1211 East Morehead Street
: CITY: Charlotte
: STATE: No. 5811407ch Carolina
: COUNTRY: USA
: ZIP: 28234
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/801,263A
: FILING DATE: 19-FEB-1997
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Sibley, Kenneth D.
: REGISTRATION NUMBER: 31,665
: REFERENCE/DOCKET NUMBER: 5470-147
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-420-2200
: TELEFAX: 919-881-3175
:
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 11703 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-801-263A-8
:
: Query Match 26.1%; Score 1303.6; DB 1; Length 11703;
: Best Local Similarity 56.1%; Pred. No. 0;
: Matches 2821; Conservative 0; Mismatches 2129; Indels 80; Gaps 16;

```


QY 437 CGTGTAGTACTTTGGAAAGTGTAGAGTATACCAAGATGTGTACGACGTCCATGCACGGA 496
DB 484 COTGAAACATGCGTCCCAATATTCCTCATCTAGAGACTGTA---TATCAACGCTCCCG 540
QY 497 CATCATCTTACCAACGAGCGCTTAAAGAGTTAGACAATTTATCTGATAGGCTTTGACA 556
DB 541 GAACTATCTATCATGAGCTATGAAAGGCGTGCAGCCCTGTACTGATTTGCTTGCACA 600
QY 557 CGACCCCTTTTATGTACAAAACATGTGAGGTTCTTACCTTACTTACAAACAGAACTGGG 616
DB 601 CCACCAAGTTCATGTTCTCGGCTATGCGAGTTGTAACCTGCTGACAAACAACTGGG 660
QY 617 CTGACGAGAGATATTGGAGCAGTAAACATTTGGCCTGGTAACTCAGATCTTTCAGAGA 676
DB 661 CCGAGCAGAAAGTCTTTGAAAGCGCTTACATCGACTTTGACGACAAAGCTGAGTAGAG 720
QY 677 GCAGGCTTGAAAACCTTCAATCTTAAAGAAAGAGGCTCCAACTTACTAATAAGATCA 736
DB 721 GTAGGACAGAAAATTTGTGATPATAGAGAAAGAGTTGAAGCCCGGCTCGCGGTTT 780
QY 737 TATTCCTGTTGTTCACAACTATCTACAGAAAGATAGATCACTGTTAGTACTGGCATC 796
DB 781 ATTCTCGTAGGATCGACACTTTATCCAGAACAGAGCCAGCTTGACAGACTGCGATC 840
QY 797 TTCCAACGCTTTCACCTTGAAAGAAAGTCTTAACTTCAAGTATGTGGAGCCATTG 856
DB 841 TTCCATCGGTGTTCATCTTGAAATGAAAGACGTGTAACCTTCCGCTGTGATPACATGG 900
QY 857 TCAGCTGTGAAGGATGATCATCAAAAAGATTAACATCAGCCAGAGACTATACGGTAAAG 916
DB 901 TGAGTTGGAAAGGCTACGTAAGTGAAGAAATCACATCTAGTCCGGGATCACGGGAGAA 960
QY 917 TTGAAACTTGGCGCTCCCAATGATCCGAGAGGTTTCTTGAGTTGCAAAAGTCAAGATA 976
DB 961 CCGTGGGATACCGCGGTACACAAATAGGAGGCGTCTTGTATGCAAAAGTTACTGACA 1020
QY 977 CGCTCGCGGCGAGAGGTTTCTTGTGTGTATACGATATGATACAGCAACCTTTGG 1036
DB 1021 CAGTAAAGAGAAAGGATATGCTTCTGTGTGACATGATCCCGGACCAATATGG 1080
QY 1037 ATCAGATGACAGGATTTCTGGCACTGACGTATGTGTGATGACGACAAAAATCTATTGG 1096
DB 1081 ATCAGATGATGATGTATATGCGCACGATATATCACTGACGATCAAAAAATCTTTGG 1140
QY 1097 TTGGCTCAACCAAGATGTGTGATGATGTAGAGCAAGAAATCTAACACAAATGC 1156
DB 1141 TTGGCTCAACCAAGATGTGTGATGATGTAGAGCAAGAAATCTAACACCAATGC 1200
QY 1157 AGAATATCTATTAACAGTGTGCGCCAGGCGTTTCCAGGTGGGCGGTGACATCGTG 1216
DB 1201 AAAATTTACTTCTGCGATCATAGCAAAAGGTTCAAGAAATGGGCTAAAGGCGCAAG 1260
QY 1217 CCGACTTGGACGACAGAAAGAACTAGGAGTGTGCGGAGCGCACTTTACTATGAGCTGT 1276
DB 1261 ATGATCTTGATTAAGAGAAAATGCTGGGTACTAGAGAAAGCAAGTTTACATGAGCTGT 1320
QY 1277 GCTGGCTTTTCAAGACCAAGAAATCAATCTTACAGAAAGCTGTGACCAACA 1336
DB 1321 TGTGGGCTTTTGCACCTAAGAAAGTACATTCGTTTATGCGCCACCTGGAACGACAGACA 1380
QY 1337 TTAAGAAAGTACCTGCGCTTTTGAATCTATTTGATTCACAGCCTTACAGCCAGGCG 1396
DB 1381 TGTAAAGATCTTCAAGCTCTTTTGAAGCTTTTCCATGTGCTGATAGAGAGACTCTTT 1440
QY 1397 TCGATATGAGGCTTCCGCGTAGCTCAAGCTGCTCTTGAACCA---ACTGTCAAAACCG 1453
DB 1441 TGCCCATGTGCGTGAAGGAGAAATTTGAAACTGTGCACTTGAACAAAGAGGAGAAAC 1500
QY 1454 CACCGGCTATTGAATGCGCGATGTGAGCATCTGCGTGGCTTTACAGCAAGAGCTGAAG 1513
DB 1501 TGCTGCAAGTCTCGAGGAAATTTATGATGAGCCAAAGGCTGCTTTTGGAGATGCTCAG 1560

QY 1514 AAGTGTGTCAGCGGAAAGATACAGAAAGCCCTGCAACCTTGCTCCCTGAAATAGAAA 1573
DB 1561 AGGAAGCCAGAGCGGAGAAAGCTCCGAGAAAGCACTTCCACATTAATGTGACAGCAAGGCA 1620
QY 1574 AAGAGCCGTAGAGCA-----GAATGACCTCATTAATGCAAGAGCGAGAGCAG 1624
DB 1621 TCGAGGCAAGCGCAGAAATGTTGTCTGCGAAGTGTGAGAGGCGCTCCAGCGGACATCGAGCAG 1680
QY 1625 GTAAGGTGAGAACACCAAGAGAGACATCAGAGTGTACAAAGTTACCAGCGGAGAGAA 1684
DB 1681 CATTAATGTAACCCCGCGGTGACGTAAAGATATATCTTCAAGCAATATGACCGTATGA 1740
QY 1685 TTGGGTCTTACGCTATCTTCAACCCAGCGGATTAATATAGTAAATACTGGCGGTGA 1744
DB 1741 TCGACACTATATGCTGTGTCTGCGCAACTCTGTGCTGAAGATTCCAAACCTCGACAG 1800
QY 1745 TCCACCAATGGCGGAAACAAATCTGTAATGACTACAAAGTATGAGGCGAGAGATCA 1804
DB 1801 CGCACCCGCTAGCAGATCAGGTTAAAGATCATTAACACACTCGGTATGATCAGGAAGTACG 1860
QY 1805 AAGTGCAGCCATACACAGGTAAGTCAATTGTACAGAAAGGAGAGCGGCTCCCTGTCAG 1864
DB 1861 CCGTGAACCATACGACGCTAAAGTACTGATGCCAGAGAGGTGCGTACCATGCGAG 1920
QY 1865 ACTTCAGGCAATGAGTGTAGAGCGCTACGATCGTTTCAACGAGAGAGGATTCGTAACA 1924
DB 1921 AATTCCTAGCACTAGTGTAGAGCGGACGCTTATGTATCAAGAAAGAGATTTGTAACC 1980
QY 1925 GATACCTGACACATCGCAATCAACGAGAGAGCGCTAAACATGACGAAAGTACTATA 1984
DB 1981 GCAAACTATACCATCTTCCATGCAATGCGCCCGCAAGATATACGAAGAGAGGACGTACA 2040
QY 1985 AGACTGTAAAGATCAGACACAGACTCAGAAATAGTCTTGATTTTGAACGCAAGAAAGT 2044
DB 2041 AGTTTAAAGAGCAGAGCTTCAAGAAACAGATAGTGTGACCTGTGACAAAGACGTT 2100
QY 2045 GTGTTAAGCAGAAAGACACGAGTCCCTGTGCTTAAACGCGTATGTGTAGTCCACAT 2104
DB 2101 GCGTTTAAAGAAAGAGAGCTTCAAGTCTGTGTCTTCTGCGGAGAACTGACCAACCTCCT 2160
QY 2105 TTCAAGATTTGCGTACAGAGTCTCAAGACAGACAGACCTGACAAAGTCCCA 2164
DB 2161 ATCATGAGACTAGCTCTGAGGAGACTGAAAGACCGGCTGCGGTCCGTAACAAAGTCGAAA 2220
QY 2165 CCAATGAGGCTATGAGATGCGCAGGTTCAAGSTAAATCTGGAATCATCAAAAGCGTGTGA 2224
DB 2221 CAATGAGGATATAGGACACCGGAGTGGCAAGTCAATTAATCAAGTCAATCTGTCA 2280
QY 2225 TTAAGAAAGATCTGTTGTGAGTGTGAGAGAAAGAAACTGGCAGAAATCATCAGGATG 2284
DB 2281 CCGCACGGGATCTTGTTCACAGCGGAAAGAAAGAAATTTGCGCAAAATGAGCGGACG 2340
QY 2285 TAAAGAGATGAGACGTATGATGTGTGCTAGAGCTGTGATTCAGTGTCTTAAATG 2344
DB 2341 TGCTAAAGACTGAGGAGTATGAGATTAAGTGAAGACAGTGAATTCGTTATGTCAAAG 2400
QY 2345 GGGTTAAGACCCCGTTTAACTCTGTACATTTGATGAGGCAATTTGCTGCCATGACGGA 2404
DB 2401 GATGCAACAAAGCGGTAAAGTGTGTACGTTGACGAAGCTTCCGTCGACAGCAGAG 2460
QY 2405 CGCTGTGCGCACTGATGCGATCGTCAAAAC---TAAAGAAAGTGTATTTGGCGGAGAC 2461
DB 2461 CACTACTTGTCTGTATGTCTATGTCAGGCGCCCGCAAGAGTATCTATGCGGAGAC 2520
QY 2462 CAATAATGCGGCTCTTTTAAATGATGTGCTGAAGTACATTTTAAAC----- 2512
DB 2521 CCAATGATGCGGATTTCTTAAATGATGAACATTAAGTACATTTCAATCAACCTGAAA 2580
QY 2513 ATGACATATGACTAGAAATGATCAATTAAGATCTTGAAGAGTGTGACAGACTGTAA 2572
DB 2581 AAGACATATGACCAAGACATTTCAAGATATCTCCGCGCTTGCACAGCGCAGTTA 2640
QY 2573 CCGCATGTGTCCACGCTCTTTACGACAAAGCAAGTAAAGCGGTTAACCATGTGTCTG 2632

Db 2641 CAGTATTTGATTCAGACTGTCATTACGATGGAAGATGAAACCAACGAAACCCGTGCAAGA 2700
Qy 2633 ATAAATATCATATGATATACCAAGAGACCAAAAGCCGCAAAAGATGATCTGATTTCTAA 2632
Db 2701 AGAACATTTAAATGATATTAACAGGGGCGCAAAAGCCGAAACCGAGGATATCATCTCTGA 2760
Qy 2693 CCTGTTTCAGAGATGGGTGAAACAGCTACAGATTGACTACAAATAATCAACGAATTCATGA 2752
Db 2761 CATTTTCCGGGGGTGGTTAGCAATTGCAATTCGACTATCCGGACATGAAGTAATGA 2820
Qy 2753 CTGGGGCTGATCCGCAAGAGACTTACGCGAAAGCGTTTATGCTGTCAAGTACAAAGTCA 2812
Db 2821 CAGCCGGGCTCTCAAGAGGCTTAACAGAAAGAGGTATGCTCCGTCCGGCAAAAAGTCA 2880
Qy 2813 ACGGAATCCACTCTACTCTGACAGCTTGTGACAGTGAAGTGTACTTACACGCACAG 2872
Db 2881 ATGAAACCCACTGTACCGCATCACTCAGAGCATGTGAAAGTGTGTCTCACCCGCACGTG 2940
Qy 2873 AAAAAGCATTTGTGGAAGACGCTAGCTGTGATCCCTGGATTAAGA CATTACAGCTA 2932
Db 2941 AGGACAGGCTAGTGTGAAAACTTGCAGGGCGACCCATGATTAAGACGCTCATTAACA 3000
Qy 2933 AATATCCGGGGATTTACCGGCTTCATTGACGACTGACGCGCAACGACGCCCATTA 2992
Db 3001 TACTTAAGGAACTTTACAGGCTACTATAGAGAGCTGGGAAGCTGAACACAAAGGAATTA 3060
Qy 2993 TGGACCGCTTCTTGATTAAGCCGAGACAGTGAATGTGTTCCAAATTAAGGTGAACGTCT 3052
Db 3061 TTGTGCAATTAACAGCCCACTCCCGTGCATTCGGTTCAGCTGCACAAAGCAACGTTT 3120
Qy 3053 GCTGGCGAAGGCTTTAGAGCCAGTCTTGGCCACGCGCAACATTGTGCTGACGAGACG 3112
Db 3121 GCTGGCGAAGGATTTGAAACCATCTAGCCACGCGCGGTATGTACTTACCGGTGCC 3180
Qy 3113 AGTGG---GAGACGTTGACCCATTCAGATGACAGAGCGTACTCACCTGAATGCGAC 3169
Db 3181 AGTGAAGGAACTGTTCCACAGTTTGGCGATGACAAACACATTCGGCCATTACGCT 3240
Qy 3170 TGAATCTTTTGAACCAAGTCTTTGAGATGACCTGGAAGTGGTATTTCCGCTC 3229
Db 3241 TAGACGTAATTTGATTAAGTTTTCGGCAGTAGACTTGACAAAGCGAATGTTTTCTAAAC 3300
Qy 3230 CTACCGTCGCACTTACTTACAG-----GGATCAGACCTGGGATA 3268
Db 3301 AGAGCATTCACATACATACATATCCCGCATTCAGAGAGCGGTAGCTATTTGGACA 3360
Qy 3269 ACTGCGCAGGAGAAACATGTATGGCTTAATTAAGAGATGACAAAGAGTTGTCAAGGC 3328
Db 3361 ACAGCCAGGAAACCCGCAAGTATGGGTACATCAAGCCATTCGCGCGAATCTCCGTA 3420
Qy 3329 GATATCCGTGATCACAAAACCGGTTGACAGAGCAGGTTAGCTGATTAAGAAATATA 3388
Db 3421 GATTTCCGGTGTCCAGCTAC---TGGGAAGGGACACAACTTGATTTGACAGCGGGA 3477
Qy 3389 CCATCAAGAACTACTCTCCAACTAATTAATGTGTTCACTTAATCGCGGTTGCCACT 3448
Db 3478 GAAACAGATTAATCTCTGACACAGATTAACCTGTCTCCGTTAACCGCAATCTTCTCAG 3537
Qy 3449 CGTTGATGTTGACCAAAAGGACAGGTTACATGATCAGCGGATTCCTATCTAAGA 3508
Db 3538 CTTAGTCCCGAGTACAAAGAGACAAACCGGCGCGGTGAAATAATTTCTTGAACAGT 3597
Qy 3509 TGAAGGCAATCTGTGTGTGATC---GGCGATCTATACAGATTCACAGGAAAGAAAG 3565
Db 3598 TCAAAACCACTCTGACTCTTGTGTATCAAGAGAAAAATGAAGCTCCCGTAAAGAA 3657
Qy 3566 TAGAGTCATGGTCCATTTGCCAC-----TAATACATCAGGTTGATCTCGATTTGG 3619
Db 3558 TCGAATGATGCCCCCATTTGGCATAGCCGTTGACAGATTAAGAACTAACCTGGCTTTCG 3717
Qy 3520 GAATACCTAGCCATGTGGTAAATATGACATTAATCTTTGTCAATGTTAGACCCCGTACA 3679

Db 3718 GGTTCGCCCGCAGGACG---GTACGACCTGTGTTTCATCAACATTGGAATAAATACA 3774
Qy 3680 GGAACATCACTAACCAAGTGGAGATGACGCTATCCACACAGCATGTAAAGTGA 3739
Db 3775 GAAACCACTTTTCACAGTGGAGAACCAATGGGAGACTTAATAAACCTTTCCGCTT 3834
Qy 3740 AGGCTGTCCACCACTTGAACACTGGCGAAACATGTGTGCTAATAGGATATGGCTTTCG 3799
Db 3835 CGGCTCTGAATTTGCTTAACCAAGAGCAACCTCGTGTGAAGTCTTATGGCTAACCCG 3894
Qy 3800 ATGCGCAACCGAGATATATCATCTCGGTGGACGCTCATTTAGTTATCCGTGCT 3859
Db 3895 ACCGCAACAGTGAAGAGTGTACCCGCTCTTGCACAAAGATTGTCAAGGTGTCACG 3954
Qy 3860 GTACGCTTAAGAACCTGCGGAAATATCTAGGTTCTTTCGTGTTCTTGGCAAGACA 3919
Db 3955 CGAGCCAGATTTGTGTTCAAGCAATACAAATATGACTGATTTTCGACAACTAACA 4014
Qy 3920 ACGGCAACCAACACATGACCAAGACAGACTCGGTGTAGTGTGACAAATCTAT-CAA 3978
Db 4015 ACAGC-CTTACAGGCAATTCACCCGCACTCATGTAATTCGATTTGTCGCTGAT 4073
Qy 3979 GGGTCAACCAAGTACGAGGAGGAGACTCCAGCTTACAGATGATCAAGTGAAT 4038
Db 4074 GAGGTTCAAGAGATGAGTGAAGCGCGCTGTCAATACGCAACCAAAAGGAGAAATTT 4133
Qy 4039 AGCAAGCGCTGACCAAGCTATGTTATGTGTCTAATAGCAAAAGTCAACAGGTTCC 4098
Db 4134 GCTGACTGTAAAGAGAGAGAGCTTTTCAACGACCAATCCGCTGGTAAACAGCGAA 4193
Qy 4099 GGAATGTGCGGTGACCTGTACCGAAATGSCCGGCTGCTTTTGAATAGACAGCAATAGCT 4158
Db 4194 GGAATGTGCGGTGACCTGTATTAACGTTGCGGACGATTTTACCGATTCAGCCACGAG 4253
Qy 4159 GTGGGACGGCTAGACT---TGTGAACAGAAACCGCTCATATCATCTGTGTGACACC 4215
Db 4254 ACAGGACCGCAAGAAAGTACTGTGTCTGAAAGAAAGTGTATCCAGCGGTGCGCT 4313
Qy 4216 AATTTTCTAAGATGCGGAAACCGAAGGCGAAGCTTAAGCTGCGAGCTGCTACATAGC 4275
Db 4314 GATTTCCGAAACACCAAGAGAGAGAGCTTGAATTTGCTTACAAACGCTTACATGCA 4373
Qy 4276 ATAGCGTCATCTGCAACGCTGAGCGGATTAACAAATATCAATACCGTACTGTCAAC 4335
Db 4374 GTGGCAACTTAATTAATGAACAATACATCAAGTGTGCTGCCATTCACATGCTATACA 4433
Qy 4336 GGCATCTATTTCTGTGCAAAAGATCGAGTATGCAATATTTGATCTCTTTTCACTGCT 4395
Db 4434 GGCATTTACCAAGCGGAAAGAACCGCTTGAAGTATCACTTAATGCTTGAACAAACGCG 4493
Qy 4396 TTGACACTACGAGTGCAGTGTCAACATATTTGTTGATTAACATAGGAGACAGG 4455
Db 4494 CTAGACGAATCTAGCGGAGCTTACATATTTGCTGTGAATTAAGAGTGAAGAAAGA 4553
Qy 4456 ATATCAGGCACTTCAACGCAAAAGAAAGCGTCAAAATTC---GATATGACAAACCA 4512
Db 4554 ATGAGCGGAGACTCAACTTAAGAGTCTGTAAACAGGCTGAAGATGAAGATATGAG 4613
Qy 4513 GTAGACTTGACTTTGTCAGGTTCAACCAACAGCTTTTGGCAGGACAAACAGTTAC 4572
Db 4614 ATCAGCATAGATTAATGATTCATACAGAGTGTGAAGGAAAGGAGATTC 4673
Qy 4573 TCGTCAATGAGGCAAGTTGATTTACCTGGAAGTACAGATTCATCAGACGCC 4632
Db 4674 AGTACTAAGAAAGAAATTTGATTTGATTTGAAAGGACCAATTTCTCATAGCGCA 4733
Qy 4633 AAGGCAATTCGCAAAATCAATGCAATGTGCCCAAAATCTGAGCTAATGACAGATT 4692
Db 4734 AAGACATGCGGAGATTAAGTCTGTCTTAAATGACAGGAAAGATTAAGAACTAG 4793
Qy 4693 TGTCTTGAATCTCTGGGAGAGATATTCAGACATCCGCTTCAATGCTCCAGTAGAGAG 4752
Db 4794 TGTGCTTACATATTTGGGTGAGACCATGGAAGCAATCCGCAAAAGTCCGCTGACCAT 4853

```

OY 4753 TCAGAGGCGCTGCTGCTCACTTCATGCTGTATTAATTAAGCTATGACGGCT 4812
    |||||
Db 4854 AACCGTGTGTAGCCCGCCAAAGTTGCCGTGCTTTGCATGTATGCCATGACGCCA 4913
    |||||
OY 4813 GAGCGGTATTAAGAGTTGGCTCTGCGAAGAAAGAACAGTTGCCGTATGCTCATCATTC 4872
    |||||
Db 4914 GAAAGGGTCCACAGACTTAGAAGCAATACGTCAAGAAGTTACAGTATGCTCTCCACC 4973
    |||||
OY 4873 CTGTGCCGAAGTACAGATCAAGGCGTGCAGAAAGCTACAGTGCAGCAAAACAGTCCCTG 4932
    |||||
Db 4974 CCCCTTCTTAAGCACAATAATTAGAATGTTCAAGAGTTCAAGTGCACGAAAGTAGTCTCTG 5033
    |||||
OY 4933 TTTTCAGGCGTGTACCAACCGGCTGTACACCCCGAAGAGTACCGGAAT 4982
    |||||
Db 5034 TTTAATCCGCACTCCCGCATTCGTTCCCGCCGTAAGTACATAGAACT 5083
    |||||

```

Search completed: November 15, 2003, 20:02:56
 Job time : 382 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:11:16 / Search time 1177 Seconds
(without alignments)
11467.451 Million cell updates/sec

Title: US-10-023-649-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 accctacaactatcgtatc.....ataattctagaacgcacc 5000

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq.19Jun03:*

- 1: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4492	89.8	11492	20	AAV74107
2	2118	42.4	11459	21	AAA49442
3	2116.4	42.3	12523	24	AAD29131
4	2113.2	42.3	12379	24	AAAD29139
5	2113.2	42.3	13584	24	AAD29141
6	2111	42.2	7479	24	AAAD29132
7	2074.8	41.5	11464	20	AAV74108
8	1565.6	31.3	11060	20	AAV78130

9	1565.6	31.3	12464	21	AAC55473
10	1564	31.3	7399	25	ACC45139
11	1564	31.3	15538	20	AAV76582
12	1562.4	31.2	11489	24	ABN86687
13	1562.4	31.2	13599	24	ABN86690
14	1554.4	31.1	11517	13	AAQ26021
15	1473.8	29.5	8010	20	AAV78129
16	1473.8	29.5	8100	20	AAV78588
17	1333.2	26.7	11663	19	AAV34475
18	1333.2	26.7	11663	19	AAV34475
19	1331.6	26.6	11663	18	AAV47668
20	1328.4	26.6	11617	21	AAA70577
21	1306.2	26.1	11626	21	AAV70607
22	1305.2	26.1	11703	21	AAC64506
23	1305.2	26.1	11703	21	AAC64507
24	1303.6	26.1	11703	19	AAV34476
25	1302.2	26.0	11717	19	AAV34474
26	1300.4	26.0	8000	25	ABX81525
27	1300.4	26.0	11740	25	ABX81526
28	1300.4	26.0	11927	20	AAV73556
29	1300.4	26.0	11927	21	AAV90388
30	1298.8	26.0	9951	20	AAV73354
31	1298.8	26.0	9951	21	AAA90386
32	1298.8	26.0	9951	22	AAV73174
33	1298.8	26.0	9951	24	ABV73174
34	1298.8	26.0	9951	24	ABN86686
35	1298.8	26.0	12110	24	ABN86689
36	1298.8	26.0	13905	20	AAV73358
37	1298.8	26.0	13905	21	AAA90390
38	1298.8	26.0	15958	25	AAV5269
39	1297.2	25.9	10524	20	AAV73355
40	1297.2	25.9	10524	21	AAA90387
41	1295.6	25.9	8000	18	AAV58321
42	1295.6	25.9	8000	20	AAV58571
43	1295.6	25.9	8000	25	ABX81524
44	1295.6	25.9	11282	20	AAZ27850
45	1295.6	25.9	11282	22	AAF84024

ALIGNMENTS

RESULT 1	ID	AAV74107 standard: CDNA, 11492 BP.	ALIGNMENTS
AAV74107	AAV74107		
AC	AAV74107;		
XX			
DT	12-APR-1999 (first entry)		
XX			
DE	Western equine encephalitis virus CDNA.		
XX			
KW	WEE virus; vaccine; ds.		
XX			
OS	Western equine encephalitis virus.		
XX			
PN	MO9853077-AL.		
XX			
PD	26-NOV-1998.		
XX			
PF	20-MAY-1998; 98MO-US10645.		
XX			
PR	16-DEC-1997; 97US-0991840.		
XX			
PR	20-MAY-1997; 97US-0047162.		
XX			
PR	24-JUL-1997; 97US-0053652.		
XX			
PA	(REED-) REED ARMY INST RES WALTER.		
XX			
PI	Cribe BJ, Oberste MS, Parker MD, Schmura SM, Smith JF;		
XX			
DR	WPI, 1999-045316/04.		
XX			
PT	New DNA encoding infectious Western or Venezuelan equine		

Destination vector
Alphavirus replicon
Plasmid pMP44 nucl
Nucleotide sequenc
Nucleotide sequenc
pSP6-SFV4 RNA tran
SemiLi Forest vir
SFV EcoRV-Spel fra
South African Arbo
South African Arbo
South African Arbo
South African Arbo
South African Arbo
Sindbis-like virus
Sindbis-like virus
Alphavirus SindCh
Alphavirus SindCh
Sindbis virus cion
Girdwood S.A.virus
Sindbis virus CDNA
Sindbis virus CDNA
Polynucleotide seq
Plasmid pSindRep5 h
Polynucleotide seq
Sindbis expression
Alphaviral vector
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Polynucleotide seq
Plasmid pRES 2U DN
DNA of expression
Polynucleotide seq
Plasmid pSindRep5 E
Sindbis virus vari
Sindbis virus CDNA
Complete CDNA sequ
Complete CDNA sequ

1607 TGCAGAGCAGAGCAGGTACGCTGAGACAC-ACGAGACACATCAGGCTGACAAGT 1665
1598 TGCAGAGCAGAGCAGGTACGCTGAGACACACACAGACACATCAGGCTGACAAGT 1657
1666 TACCCAGCGC-AAGAGAGATGGGCTTTA-CGCTACTTTCACCCAGCGGATTT-G 1722
1658 TACCCAGGAGAGAGAGATGGGCTTTATCCCTACTTACCCAGCGGCTTTTAN 1717
1723 AATAGTGAAGAACTGCGCTGATTCACCCATGGGGGAAACAAGTCTGTAATGACTC 1782
1718 AATAGTGAAGAACTGCGCTGATTCACCCATGGGGGAAACAAGTCTGTAATGACTC 1777
1783 AAGGTAGGGCAGGAGATACAAAGTCAGCCATACACGCTAGGCTATTGTA-CCAGA 1841
1778 AAGGACGGGCGGGGATACAAAGTCAGCCATACACGCTAGGCTATTGTAACCCAGA 1837
1842 AAGGACGGC-GGTCCTGTTCAAGACTTCAGGCAATTGATGAGAGCGCTACGATCGTT 1900
1838 AAGGACGGGCGGCTCCCTGTTCAAGACTTCAGGCAATTGATGAGAGCGCTACGATCGTT 1897
1901 TCAAGAGAGGAGTTCGTA------CAGATACCTGCACCCATCCGAACTA 1948
1898 TCAAGAGAGGAGTTCGTA-CAAGATTTTGCACCCATCCGAACTTTTCAACTA 1957
1949 ACGAGAGCGCTTAACACTGACGAAGATCTATAGACTGTAAAGACTCAGAGACAG 2008
1958 TAGTAGTCGCTATTACACTGACGAAGATCTATAGATTTTAAAGACTCAGAGACAG 2017
2009 ACTCAAGATAGCTCTTCATATTTGACGACGAAGTGTGTTAAGCGAAGACGAGCTC 2068
2018 ACTCAAGATAGCTCTTCATATTTGACGACGAAGTGTGTTAAGCGAAGACGAGCTC 2077
2069 CCTTGCTCTTACCGGCTGATCTGTAAGATCCACATTTTACAGATTTGGTACGAGATC 2128
2078 CCTTGCTCTTACCGGCTGATCTGTAAGATCCACATTTTACAGATTTGGTACGAGATC 2137
2129 TCAAGACACGACGACGACCTCACAAGTCCCAACCATCGAGCTATGAGAGTCCAG 2188
2138 TCAAGACACGACGACGACCTCACAAGTCCCAACCATCGAGCTATGAGAGTCCAG 2197
2189 GTTCAGTAAATCTGGAATCATCAAAAGCGCTGTGATCAAGAAAGTCTGTTGATG 2248
2198 GTTCAGTAAATCTGGAATCATCAAAAGCGCTGTGATCAAGAAAGTCTGTTGATG 2257
2249 CGAAGAGAGAACTGCGGAGAAATCATAGGAGTGTAAAGAGTGAAGCTATGAGATG 2308
2258 CGAAGAGAGAACTGCGGAGAAATCATAGGAGTGTAAAGAGTGAAGCTATGAGATG 2317
2309 TTGCTGCTAGAGCTGTCATAGTCTCTAAATGGGGTTAAGCACCCGTTAACATC 2368
2318 TTGCTGCTAGAGCTGTCATAGTCTCTAAATGGGGTTAAGCACCCGTTAACATC 2377
2369 TGTACATTGATGAGCATTTGCTGCTGATGACGAGGACGCTGCTGAGACTGATTTGCAT 2428
2378 TGTACATTGATGAGCATTTGCTGCTGATGACGAGGACGCTGCTGAGACTGATTTGCAT 2437
2429 TCAAACTTAAGAAAGTGTGATGCTGGGAGACCCAAACAAATGCGGCTTTTAAACATGA 2488
2438 TCAAACTTAAGAAAGTGTGATGCTGGGAGACCCAAACAAATGCGGCTTTTAAACATGA 2497
2489 TGTGCTGAAAGTACATTTTAAACATGACATGACATGACATGACATGACATGACATG 2548
2498 TGTGCTGAAAGTACATTTTAAACATGACATGACATGACATGACATGACATGACATG 2557
2549 CTAGAGAGTGCACACAGACTGTATACCGGCACTTGTCCAGCGCTTTCTACGAAGCGAA 2608
2558 CTAGAGAGTGCACACAGACTGTATACCGGCACTTGTCCAGCGCTTTCTACGAAGCGAA 2617
2609 TGAAGACGTTAAACCATGTGCTGATTAATCATCATAGATACACAGGAGACCAAAAGC 2668
2618 TGAAGACGTTAAACCATGTGCTGATTAATCATCATAGATACACAGGAGACCAAAAGC 2677

2669 CGCACAAAGATGATCTGATTTAACTGTTTCAGAGATGGGTGAACAAGCTACAGATTG 2728
2678 CGCACAAAGATGATCTGATTTAACTGTTTCAGAGATGGGTGAACAAGCTACAGATTG 2737
2729 ACTACAAATTCAGAAATCATGACTGCGGCTGCATCCGAAGGACTTACGCGAAAGCG 2788
2738 ACTATMAAATCATGAAATCATGACTGCGGCTGCATCCGAAGGACTTACGCGAAAGCG 2797
2789 TTTATGCTGACGATACAAAGTCAAGGAAATCCACTGATCCGAGACTTCTGAGCAG 2848
2788 TTTATGCTGACGATACAAAGTCAAGGAAATCCACTGATCCGAGACTTCTGAGCAG 2857
2849 TGAAGTGTACTTACACGACAGAAAAACGATTTGTGGAAGAAGCTAGCTGTGATC 2908
2858 TGAAGTGTACTTACACGACAGAAAAACGATTTGTGGAAGAAGCTAGCTGTGATC 2917
2909 CCTGATTAAGACACTTACAGCTAAATATCCCGGGGATTTCAAGGCTTCAATTGACGACT 2968
2918 CCTGATTAAGACACTTACAGCTAAATATCCCGGGGATTTCAAGGCTTCAATTGACGACT 2977
2969 GGCAGCGGAAACAGAGCGCATTTATGACAGCGCTTGTATAGCGGAGACAGCTGATG 3028
2978 GGCAGCGGAAACAGAGCGCATTTATGACAGCGCTTGTATAGCGGAGACAGCTGATG 3037
3029 TGTTCAGAAATAGGTGAAGCTCTGCTGGGAGAGCTTTAGAGCGAGCTTTGCGCACGG 3088
3038 TGTTCAGAAATAGGTGAAGCTCTGCTGGGAGAGCTTTAGAGCGAGCTTTGCGCACGG 3097
3089 CCAACATTGTGCTGACGAGACAGAGCTGGAGAGCTTGCAACCATTCACAGTACAGAG 3148
3098 CCAACATTGTGCTGACGAGACAGAGCTGGAGAGCTTGCAACCATTCACAGTACAGAG 3157
3149 GCTACTACCTGAAATGGGACCTGAACTTTTGAACAGAGTCTTTGAGAGTACCTG 3208
3158 GCTACTACCTGAAATGGGACCTGAACTTTTGAACAGAGTCTTTGAGAGTACCTG 3217
3209 ACAGTGGGTTATTTCCGCTCTTACCGCTCGACCTTACTTACAGGAGTACGACTGGATA 3268
3218 ACAGTGGGTTATTTCCGCTCTTACCGCTCGACCTTACTTACAGGAGTACGACTGGATA 3277
3269 ACTGCCAGGAGAGACATGATGAGGCTTTAATGAGAGTACGAAGAGTGTGACAGGC 3328
3278 ACTGCCAGGAGAGACATGATGAGGCTTTAATGAGAGTACGAAGAGTGTGACAGGC 3337
3329 GATATCCGCTGATACCAAAACCGGTTGACAGAGAGGCTGATTAAGAAATTAATA 3388
3338 GATATCCGCTGATACCAAAACCGGTTGACAGAGAGGCTGATTAAGAAATTAATA 3397
3389 CCATCAAGGACTACTCTCCAACAATTAATGTTGCTTCAATTAATGCGCGTTGCCCACT 3448
3398 CCATCAAGGACTACTCTCCAACAATTAATGTTGCTTCAATTAATGCGCGTTGCCCACT 3457
3449 CGTTGATGTTGACCAAAAGAGACAGGTTACACTGATCAGAGCGGATTCCTATCTAAGA 3508
3458 CGTTGATGTTGACCAAAAGAGAGGTTACACTGATCAGAGCGGATTCCTATCTAAGA 3517
3509 TGAAGGAGAAATCTGTGTTGATGCGGATCCATACAGCTTCCAGGAGAAAGATG 3568
3518 TGAAGGAGAAATCTGTGTTGATGCGGATCCATACAGCTTCCAGGAGAAAGATG 3577
3569 AGTCCATGGGTCATATGGCCACTAATACATCAGGTTGATCTGATTTGGGAATACCTA 3628
3578 AGTCCATGGGTCATATGGCCACTAATACATCAGGTTGATCTGATTTGGGAATACCTA 3637
3629 GCCATGTCGGTAAATATGACATTAATCTTGTCAATGTTAGGACCCCGTACAGAAACATC 3688
3638 GCCATGTCGGTAAATATGACATTAATCTTGTCAATGTTAGGACCCCGTACAGAAACATC 3697
3689 ACTACCAACATGTCGAGAGATACGCTATCCACACAGATGCTAAACGTTAAGCTGTCC 3748
3698 ACTACCAACATGTCGAGAGATACGCTATCCACACAGATGCTAAACGTTAAGCTGTCC 3757
3749 ACACCTGAACATGTCGAGAGATGCTATAGGTTAGGCTTCTGATCGCGCA 3808

```
Db 3758 ACCACCTGGAACACTGGGGGAACATGTGTGGCCATAGGTTAGCATTTGCTGATGCGGAA 3817
Qy 3809 CCGAGAAATATCATCTACTGCGGTGGCAAGCTCATTTAGTTTACCCTGTCTGTACGCTTA 3868
Db 3818 CCGGAAATATCATCTACTGCGGTGGCTCGCTCATTTAGTTTACCCTGTCTGTACGCTTA 3877
Qy 3869 AGAACATCGCCGAAAATCTAGAGTTCCTCTGTGTTCTTGGGAAGGACAAACGCAACC 3928
Db 3878 AGAACATCGCCGAAAATCTAGAGTTCCTCTGTGTTCTTGGGAAGGACAAACGCAACC 3937
Qy 3929 ACACACATGACACAGACAGACTCGGTGTAGTCTTGAACAATCATCAAGGTCACCA 3988
Db 3938 ACACACATGACACAGACAGACTCGGTGTAGTCTTGAACAATCATCAAGGTCACCA 3997
Qy 3989 GGTACGAGGACGAGGAGAGCTCCAGGCTACAGAGTATCAGAGGTGACATTACCAAGCG 4048
Db 3998 GGTACGAGGACGAGGAGAGCTCCAGGCTACAGAGTATCAGAGGTGACATTACCAAGCG 4057
Qy 4049 CTGACCAAGCTATGTGTTAATGCTGTCTATAGCAAAAGGTCAACGAGTTCGAGTGGCG 4108
Db 4058 CTGACCAAGCTATGTGTTAATGCTGTCTATAGCAAAAGGTCAACGAGTTCGAGTGGCG 4117
Qy 4109 GTGCACTGTACCGAAAATGCGCGCTGCTTTTGAATAGACACCAATAGCTGTCCGGAACG 4168
Db 4118 GTGCACTGTACCGAAAATGCGCGCTGCTTTTGAATAGACACCAATAGCTGTCCGGAACG 4177
Qy 4169 CTAGACTTTGAGACGACGACCGCTCATCATCATCTGTATGAGACCCAAATTTTCTAAGA 4228
Db 4178 CTAGACTTTGAGACGACGACCGCTCATCATCATCTGTATGAGACCCAAATTTTCTAAGA 4237
Qy 4229 TGCCCGAACCGGAGGCGACCTTAAGCTGCGAGCGCTTACATAGCATATAGCATTCGTCATCG 4288
Db 4238 TGCCCGAACCGGAGGCGACCTTAAGCTGCGAGCGCTTACATAGCATATAGCATTCGTCATCG 4297
Qy 4289 TCACACGCTGAGCGGATTACAAAATATATCATAGTACCGCTACTGTCAACCGGCACTTATCTG 4348
Db 4298 TCACACGCTGAGCGGATTACTAAATATATAGTACCGCTACTGTCAACCGGCACTTATCTG 4357
Qy 4349 GTGCGAAGATGAGATGATGCAATCATCTGCTGTGCTGCTTGCACACTACG 4408
Db 4358 GTGCGAAGATGAGATGATGCAATCATCTGCTGTGCTGCTTGCACACTACG 4417
Qy 4409 ATGCGAGTGTACCATATATGCTTGGATTAACAATGGGAGACCAAGATATTCAGAGGCA 4468
Db 4418 ATGCGAGTGTACCATATATGCTTGGATTAACAATGGGAGACCAAGATATTCAGAGGCA 4477
Qy 4469 TTCAACCGCAAGAAAGCTGCAAAATTTCTG--GATGATGACAGGCAAGTATGACATTGACT 4525
Db 4478 TTCAACCGCAAGAAAGCTGCAAAATTTCTG--GATGATGACAGGCAAGTATGACATTGACT 4537
Qy 4526 TGGTCAAGGTCACCCCAACAGCTTTTGGGACGACGACCAAGTTATCTCCGTCAATGAGG 4585
Db 4538 TGGTCAAGGTCACCCCAACAGCTTTTGGGACGACGACCAAGTTATCTCCGTCAATGAGG 4597
Qy 4586 GCAAGTTGTATTCATCTGAGAGGTACAGATTCACATCCAGCGCAAGAACATTTGGCG 4645
Db 4598 GCAAGTTGTATTCATCTGAGAGGTACAGATTCACATCCAGCGCAAGAACATTTGGCG 4657
Qy 4646 AAATTCATGCAATGTGGCCCAACAAATCTGAGGCTTAATGAGCAGATTTGCTTGTATCATCC 4705
Db 4658 AAATTCATGCAATGTGGCCCAACAAATCTGAGGCTTAATGAGCAGATTTGCTTGTATCATCC 4717
Qy 4706 TGGGGGAGAGTATGTCCAGCATCCGCTCAAAATGCCAGTATGAGAGATCAGAGGCTGTG 4765
Db 4718 TGGGGGAGAGTATGTCCAGCATCCGCTCAAAATGCCAGTATGAGAGATCAGAGGCTGTG 4777
Qy 4766 CTCAACCTCACAACCTTCATAGCCGTGTAAATTAAGCATAGCGGCTGAGCGGATATCA 4825
Db 4778 CTCAACCTCACAACCTTCATAGCCGTGTAAATTAAGCATAGCGGCTGAGCGGATATCA 4837
Qy 4826 GGTTCGCTCTGCGAAGAAAGACAGTTCGCGTATGCTCATCTCTGTTGCCGAGT 4885
```

```
Db 4838 GGTTCGCTCTGCGAAGAAAGACAGTTCGCGGTATGCTCATCTCCGTCGCGAAGT 4897
Qy 4886 ACAGATACACAGCGCTGCAGAGCTACAGTGCAGCAACAGATCCGTTTTCAGGCGTGC 4945
Db 4898 ACAGATACACAGCGCTGCAGAGCTACAGTGCAGCAACAGATCCGTTTTCAGGCGTGC 4957
Qy 4946 TACCACCGGCTGTACACCCGAGAGTACCGGAAATATTTCTAAGAAACCCACCC 5000
Db 4958 TACCACCGGCTGTACACCCGAGAGTACCGGAAATATTTCTAAGAAACCCACCC 5012

RESULT 2
AAA49442
ID AAA49442 standard; DNA; 11459 BP.
XX
AC AAA49442;
XX
DT 25-SEP-2000 (first entry)
XX
DE Venezuelan equine encephalitis virus (VEE) wild type sequence.
XX
KM Venezuelan equine encephalitis virus; VEE; attenuation; vaccine;
KM replicon; heterologous gene expression; capsid; structural protein;
KM glycoprotein; immunogen; ds.
XX
OS Venezuelan equine encephalitis.
XX
FN WO200034445-A2.
XX
PD 15-JUN-2000.
XX
PE 07-DEC-1999; 99WO-US29041.
XX
PR 07-DEC-1998; 98US-0111330.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Pushko P, Parker MD, Smith JF, Crise BJ;
XX
DR WPI; 2000-423405/36.
XX
PT Attenuated Venezuelan equine encephalitis virus useful as a vaccine
PT contains a rearrangement in genome such that its structural
PT glycoproteins precede the capsid gene
XX
PS Disclosure; Page 37-43; 43pp; English.
XX
XX
CC An attenuated Venezuelan equine encephalitis (VEE) has been described
CC in which the virus undergoes a rearrangement of its genome such that
CC the structural glycoproteins precede the capsid gene. The resulting
CC attenuated virus is useful as a live VEE vaccine for human and
CC veterinary use. The attenuated virus can also be used in a replicon
CC expression system. The rearranged VEE replicon can be used to express
CC heterologous genes of interest in cells as well as a means for
CC expressing antigens or immunogenic proteins and peptides of interest.
CC This rearranged sequence appears in the attenuated virus designated
CC RV1.1.1.
XX
SQ Sequence 11459 BP; 3242 A; 2804 C; 2897 G; 2516 T; 0 other;
XX

Query Match 42.4%; Score 2118; DB 21; Length 11459;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3219; Conservative 0; Mismatches 1745; Indels 24; Gaps 3;
```

Db 150 GACCAATGCTAATGCGAGAGCGTTTTCGATCTGGCTTCAAACTCATCGAAACGAGG 209
Qy 190 GACCGGACCAAGTATCTTGGACATTTGGAGTGGCCCGGTGACAGATGCAATTCAT 249
Db 210 GACCCATCCGACACATCTTGAATTTGAATGGGCGCCCGGCAAAATGTAATCTTAAG 269
Qy 250 CACCCATCATTTGATCTGCGCTTATGATTAAGCGCTGAAGACCCGGACAGATACAAAG 309
Db 270 CACAGTATCATTTGATCTGCTCGATGAGATGGGAAATCCGACAGATTTGATTAAG 329
Qy 310 TATGCAAAAGACTTAAAGAAAGT-----GACATTAACGACAAAGATAGCTCTTAAG 363
Db 330 TATGCAATTAAGCTAAAGAAAGTAAAGAAATTAATGATTAAGAAATTTGACAAAGAA 389
Qy 364 GCGGAGACTGCTGGAAGTCAATGTCACACAGAGCGAGATCCATCTCTGATGATG 423
Db 390 ATGAAGAGCTGCGCGCTCATGAGCGACCTTGACCTGAAACTGAGACTATGTCCTC 449
Qy 424 CACACAGACGCGACGCTAGTACTTTGGAAAGTGAAGATATACCAAGATGTAACCA 483
Db 450 CACGACGAGAGCTGCTGCTGCTACGAAAGGCAAGTCGCTTTACCAAGATTTAACC 509
Qy 484 GTCCATGCAACGACATCAATCTTACCAAGCGCTTAAAGAGTTAGACAAATTTACTGG 543
Db 510 GTTGACGAGCCGACAGTCTCTATCACCAAGCCAAATTAAGGAGTTAGAGTCGCTACTGG 569
Qy 544 ATAGCTTTGACAGACCCCTTTTATGTAACAAACATGGCAGGTTCCATCCCTACTAC 603
Db 570 ATAGCTTTGACAGACCCCTTTTATGTAACAAACATGGCAGGTTCCATCCCTACTAC 629
Qy 604 AACAGCACTGGGCTGACGAGAGATTTGAAAGACGTAATTCATTCGCTGCTTAACCA 663
Db 630 TCTACCACTGGGCGCGAGCAACCGTGTAAAGGCTCTGTAATAGGCTTAAGCACTCT 689
Qy 664 GATCTTCAGAGAGCAGGCTTGAAGAACTCTCAATCTTTAGAGAAAGAGCTCCACCT 723
Db 690 GACGTTATGAGCGGTCAAGTAGAGGATGTCATTTCTTAAGAAAGATTTGAAACCA 749
Qy 724 ACTAATTAAGATATTTCTCGGTTGGTTCAACATCTTACAGAAAGATGATCACTGTTA 783
Db 750 TCCAACTAATGTTCTATTTCTGTTGCTGCTGACATCTTACCAAGAAAGGAGCTTACTG 809
Qy 784 CGTACCTGGCATCTTCCAAACGTTGTCACCTGAAAGGAAAGTCTTAATCTTCAGAGTGA 843
Db 810 AGGAGCTGGACCTCGCTGTATTTCACTTACGCGACAGCAAAATTTACAGATGTCG 869
Qy 844 TGTGGACCATTTGTCAGCTGTGAAGGATACGTCATCAAAAGATTAACATCAGCCAGGA 903
Db 870 TGTGACATATATGTTAGTTGCGAGGATACGTCGTTAAAGAAATGCTATCAGTCAGGC 929
Qy 904 CTATACGGTAAAGTTGAGAACTTGGCGTCCACAAATGATCCGAGGCTTTCTTGAGTTGC 963
Db 930 CTGATGGAAGCCCTTCAAGGCTATCTGCTACGATGCAACCGGAGAGGATCTTGTGCTGC 989
Qy 964 AAAGTCACAGATAGCTGCGCGGCGAGAGGGTTCTTTTGGCTGATGATGATGATCA 1023
Db 990 AAAGTCACAGATAGCTGCGCGGAGAGGGTCTCTTTTCCGCTGACGATGTGCA 1049
Qy 1024 GCCACACTTTGGCATCAGATGACAGGATTTCTGGCACTGACGTTAGTGTGATGACCA 1083
Db 1050 GCTACATTTGTGACCAATGATGCTGCACTGGCAAGATGTCAGTGGGAGACGACGG 1109
Qy 1084 CAAAACTATTTGTTGGGCTCAACCAAGGATTTGCTCAATGTAAGACGCAAGAAAT 1143
Db 1110 CAAAACTGCTGTTGGGCTCAACCAAGGATTTGCTCAATGTAAGACGCAAGAAAT 1169
Qy 1144 ACTACACAAATGATCTATTTACCAAGTGTGCGCAGGCTTTTCCAGTGGGCG 1203
Db 1170 ACCAATACATTAATAATTTACTTTTCCGCTAGTGGCCAGGCAATTTGCTAGTGGGCA 1229
Qy 1204 CGTGAACATCGTGCAGACTTGGACGACGAAAGAACTAGGGGTGCGGAGCGCACTCTT 1263

Db 1220 AAGGAATTAAGAAATCAAGATCAAGAAATGAAGGCCATAGGATCAAGATGACATTA 1289
Qy 1264 ACTATGGCTGCTGCTGGGCTTTTCAAGACCAAGAAATACATCCATCACAAGAGCT 1323
Db 1290 GTCATGGGCTTTGTTGGCTTTTGAAGGCAAGATTAACATCTATTTATTAAGGCGCCG 1349
Qy 1324 GTTACGCAAAATTTAAGAAATGATCTGCGCTTTTGAAGTCAATTTGATTTCAAGCCTT 1383
Db 1350 GATACCAAAACATCATTAAGATGAAGAGCATTTCCACTATTTGCTGCTGCCAGATTA 1409
Qy 1384 ACCAGCCAGGAGCTGATATGAGCTTCCGCTGAGGCTCAAGCTGCTCTTGAACCAACT 1443
Db 1410 GGCAGTAAACATTTGAGATGAGGCTGAGAACCAAGAAATGATGAGAGAGAC 1469
Qy 1444 GTCAAAACCGGACCGGCTATTTACATGSCCGATGAGGACATCTGCGGCTTAACAGCA 1503
Db 1470 AAGAGCGCTACCTCTCTATTAACCGCCAGAGACGTAACAAGAGCTTAAGTGGCAGCT 1529
Qy 1504 GAAGCTGAAGAAATGAGCTGACGCGAAGAGATCAAGAAAGCCCTGACACCTTGTCTCCT 1563
Db 1530 GAGGCTTAAGAGAGTGCCTGAAGCCGAGAGATTTGGCGCAGCTTACCACTTTGGCAGCT 1589
Qy 1564 GAATTAAGAAAGAGACCTGTAAGGCGAGAGTGAATCTCAATTAAGCAAGGCGAGACA 1623
Db 1590 GATGTTGAGAGCCCACTCTGAAAGCCGATGTCGACTTGAATGTTAACAAGAGGCTGGGCGC 1649
Qy 1624 GGTAGCGTGAAGACCAACGAGAGACATCAAGGCTGAAGTTCACCGAGGAGAGAGAG 1683
Db 1650 GACTCAGTGAAGACACCTGCTGGCTTGAATTAAGTTACACTAGCTGCGGAGAGACAG 1709
Qy 1684 ATGGGCTTTACGCTATCTTCAACCCGAGCGGATTTGAATGTAAGAAATCTGGCTGT 1743
Db 1710 ATCGCTTTTACGCTGCTGCTTCTTCCGAGGCTGATCAAGAGTAAATTTACTTTC 1769
Qy 1744 ATCCACCATTTGGCGGAACAGTACTGTAATGATCAACAAGAGTAAAGGCGAGAGATAC 1803
Db 1770 ATCCACCTCTGCTGGAACAGTCAATGATTAACACTCTGCGCGAAAGGCGGTTAT 1829
Qy 1804 AAAGTCGAGCCATCAAGGTAAGTCAATTTACAGAAAGGAGCGGCTCCGTTCAA 1863
Db 1830 GCGGTGAACCATCAAGTGAAGTGTGTGCGAGGAGACATGAATCCGCTCAG 1889
Qy 1864 GACTTCCAGGATTTGATGAGAGCGCTACGATCGTTTCAACGAGAGGAGTCTTAAC 1923
Db 1890 GACTTTCAGCTCTGAGTGAAGTGTCCACATTTGTGTAACAGAACGATGCTTAAC 1949
Qy 1924 AGATACCTGACCAATGCAATCAACGAGAGCGCTAAACACTGACGAAGATGATAT 1983
Db 1950 AGTACCTGACCAATGTCACACATGAGAGAGGCTGAACACTGATGAAGAAATATTAC 2009
Qy 1984 AAGACTGTAAAGCTCAGAGACACAGATCAAGATTAACGTTCTGATATTGACGACGAAAG 2043
Db 2010 AAAACTGTCAAGCCCAAGGACACAGACGCAATTAACGTAAGATGACAGAAACAG 2069
Qy 2044 TGTGTTAAGCAGAAAGACGAGGTCCCTTGTGCTCAACCGGATCTGATGATCAACA 2103
Db 2070 TCGCTCAAGAAAGATTAAGTACTGAGGCTTACAGGCGAGCTGTGATCTTCC 2129
Qy 2104 TTTCAAGATTTGCTGACGAGTCTCAAGACAGACCGAGACCTCACAAATCTCCA 2163
Db 2130 TTCCATGAATTCGCTTACGAGATCTGAGAACAGACGAGCGCTCTTACCAAGTACA 2189
Qy 2164 ACCATCGAGTCTATGAGTGCAGGTTCAAGTAAATTTGAAATTAACAAAGCGCTGTG 2223
Db 2190 ACCATGAGGCTGTATGCGGTCCAGAGTCAAGTCTGCAATTAATAAGCGCAATC 2249
Qy 2224 ACTAAGAAAGATCTGCTTGTGATGTCGAAGAGAAACTGCGCAAAATCATCAGGAT 2283
Db 2250 ACCAATAAGATCTATGAGTGAAGCCCAAGAAAGAAATCTGCAAGAAATTAAGGAC 2309
Qy 2284 GTAAGAGATGAGACGATGATGATGCTGATGAGCTGATGATGATGATGATGATGAT 2343
Db 2310 GTCAAGAAATGAAGGCTGAGCGTCAATGTCAGAACTGTGACTGATGAT 2369

2344 GGGGTTAAGACCCCGTTAAACATCTGTACATTTGTAGGAGATTGCGTCATGCGAGG 2403
D 2370 GGATGCAAAACCCCGTAGAACCTGTATATGACGAGCTTTTGGCTGTGACAGAGGT 2429
Q 2404 ACGCTGTGGACATGATTTGCGATCGTCAACCTAAGAAAGTGTATTTGTGGGGACCA 2463
D 2430 ACTTCAGAGGCGTCATAGCATTATAGACTTAAGAAAAAGCAGTGTCTGGGGGATCCC 2489
Q 2464 AAAAATGCGGCTTCTTAAATGATGTGCTGAAAGTACATTTTAAACATGACATATGC 2523
D 2490 AAAAAGCGGGTTTTTTTAAATGATGTGCTGAAAGTACATTTTAAACAGAGATTTGC 2549
Q 2524 ACTGAAGTACATTAAGGATCTCTAGAGGTGACACAGACCTGTAAACCGCATGTC 2583
D 2550 ACAAGAGTCTTCCACAAAAGCATCTTCGCGTTGACATTAATCTGTGACTTCGGGTGC 2609
Q 2584 TCACGCTCTTCTACGACAAAGCAGATGAGAAGCGTTAACCCATGTGCTGATTAATCATC 2643
D 2610 TCACCTTGTTTTACGACAAAATATGAGAGAGATCCGAAAGAGACTAAGATTGTG 2669
Q 2644 ATGATATCACAGGACCAACAAAGCCGACAAAGATGATCTGATTTTACCTGTTTACA 2703
D 2670 ATTGACACTACCGGACGATCCAAACCTAAGCAGGACGATCTCATTTCTCATCTGTTTCAGA 2729
Q 2704 GGATGGGTGAACAGCTACAGATTGACCTACAAAATACGAAATCATGACTGCGGCTGCA 2763
D 2730 GGGTGGGTGAAGAGTTGCAATATGATTAACAAAGGCAACGAAATATATGACGAGCTGCC 2789
Q 2764 TCGCAGAGACTTACGCGAAAGCGTTTATGCTGTAGGTACAAAGTCAACGAGATCCA 2823
D 2790 TCTCAAGGCGTACCCGTTAAAGTGTGATGTCCTGTGCTCAAGGTGAATGAAATCTT 2849
Q 2824 CTCTACTGCGACATTTCTGAGCAGTGAAGCTGTTACTTACACGACAGAAAAAGCATT 2883
D 2850 CTGTACGACCCACCTCAGAAACATGTGAACCTCTACTGACCCGACGAGAGACCGCATC 2909
Q 2884 GTCTGAGAGAGCTAGCTGAGTGTATCCCGATTAAGACATTAAGCTAAATATCCGGG 2943
D 2910 GTGTGAGAAACACTAGCCGCGGACCCATGATTAAGAACTGATCTGCCAATGATCCCTGGG 2969
Q 2944 GATTTCACGCTTATTTGAGCAGTGTGAGCGGACGCAACGACCATTAATGCGCGTT 3003
D 2970 AATTTCATGCGCAGATATGAGAGTGTGAGCAGAGCATGATGATGATGAGGACATC 3029
Q 3004 CTGTATAGCCGCGACAGCTGATGTGTTCAGATTAAGGTGAAGCTGTGCTGGCGAAG 3063
D 3030 TTGAGAGAGCCGGAACCTTACCGAGCTTTCAGAAATAGGCAAACTGTGTTGGGCCAAG 3089
Q 3064 GCTTTAGAGCAGCTTGGGACACGCGCAACATTTGGCTGACGAGACAGCTGGGAGAGC 3123
D 3090 GCTTTAGGCGCGGTGCTGAAGACGCGTGGCATTAACATGACCATGAAATGGAACACT 3149
Q 3124 TTGACACCATTTCAAGCATGACAGAGCGTACTCACTGAATAGCACTGAATCTTTTTCG 3183
D 3150 GTGATTAATTTTGAAGCGAACAAAGCTCCTCAGAGAGATGATTTGAACCAACTATGC 3209
Q 3184 ACCAGGTTCTTTGAGATGACCTGAGCAGTGGTTATTTTCCGCTCTACCGTGCACCT 3243
D 3210 GTGAGGTTCTTTGAGCTGATCTGAGACTCCGCTCATTTTCTGACCCACGTTCCGTTA 3269
Q 3244 ACTTACAGGGAATGACGCTGGGATTAACGCGCAGGGAAGACATGTAATGGGCTTAATGA 3303
D 3270 TCCATTAGAAATATCACTGGGATTAACCTCCGTCGCTTAACATGTAAGGCGCTGAATAA 3329
Q 3304 GAGGTAGCAAAAGAGTTGTCAACGCGGATATCCGTGATCACAAAAAGCGGTGACAGGC 3363
D 3330 GAATGTGTCGTGAGCTCTCTGAGAGTACCAACAACCTGCTGGGAGGTTGCCACTGGA 3389
Q 3364 AGGAGTGTATATATGAGATATATCAATCAAGGACTACTCTCAACAATTAATGTGTT 3423
D 3390 AGAGTCTATGACATGAACTGTGTAACCTGCGCAATATGATCGCGCATTAACCTAGTA 3449

Q 3424 CCATTAATCCCGGTTGCCCACTCGTGTATCGTTGACCAAAAGACAGGGTACAACT 3483
D 3450 CTTGTAAACAAAGACTGCTCATGCTTTAGTCTTCCACCTATATGAAACCCACAGAGT 3509
Q 3484 GATCAGAGGAAATCTTATCTTAAGATGAAGGCAAAATCTGTGTTGTGATGCGGATCT 3543
D 3510 GACTTTTCTTCAATTCGACAGCAATATGAAAGGCGAATCTGCTGCGGCGGAAAG 3569
Q 3544 ATCAGCATTCAGGGAAGAAATGATAGTCCATGGTCCATTTGCCCATTAATACATAGG 3603
D 3570 TTGTCCGTCCAGGCAAAATGTTGATGCTGTGTGACAGCGGCTGAGGCTACTTACGA 3629
Q 3604 TGTGATCTGATTTGGGAATACCTAGCCATGCGGTAAATATGACATTAATCTTTGTCAAT 3663
D 3630 GCTCGGCTGATTTAGGCAATCCAGGTGATGTGCCAAATATGACATTAATTTGTTAAT 3689
Q 3664 GTTAGAACCCCGTACAGAAACATCACTACCAACAGTGCAGAGATCAGCTATCACAC 3723
D 3690 GTGAGACCCCATATTAATACATCACTATCAGCAGTGAAGACATGCCATTAACCTT 3749
Q 3724 AGCATGCTAAGTGTAGGCTGTCCACACCTGAACATGCGGGAACATGTTGGCTATA 3783
D 3750 AGCATGTTGACCAAGAAAGCTGTGTGATCTGAATCCGCGGAACCTGTGACATA 3809
Q 3784 GGGTATGGGCTTGTGATTCGCGCAACCGAATATCATCTGCGGTGGACAGCTCATTT 3843
D 3810 GGTATGTATACGTGACAGGCGCAGGAAAGATCATTTGGTGTATAGCGGCACTTGC 3869
Q 3844 AGGTTTACCCGTGTCTGTAGCTTACAGCACTGCGGAAATATCTGAGTTCTTCTGTCG 3903
D 3870 AAGTTTCCGCGGTATCAACCGAAATCCTCACTTGAAGACGGAAGTTCTGTTGTA 3929
Q 3904 TTCTTGGCAAGACCAACCGGCAACACATGACACAGACATCTCGGTGTAGTCTT 3963
D 3930 TTCAATGGGTACATTCGCAAGCCCGGTAGCACAATCTTACCAAGTTTCAATCAACCTTG 3989
Q 3964 GACAACATCTATCAAGGCTCAACAGGTAAGAGCAGGAGAGCTCAGGTAACAGATG 4023
D 3990 ACCAACTTTATACAGTTTCAAGCTCAGAGACGAGAGCGGATGACCTCATATCATGTG 4049
Q 4024 ATCAGAGTACATTTAGCAAGAGCGTACCAAGCTATCGTTAATGCTGTAATAGCAAA 4083
D 4050 GTGCGAAGGGAATATGTCACAGGCAACGAAAGAGTGTATTAATCTGCTTACAGCAAA 4109
Q 4084 GGTCAACGAGTTCCGAGGTGTGGGTCAGTGTACCGAAGAAATGCGCGGCTTTGAT 4143
D 4110 GGAACACTGCGGAGGAGGTGTGGAGCGCTGTATAGAAAGTTCCCGGAAAGCTTGAT 4169
Q 4144 AGACAGCAATAGCTGTGCGGACGCGTGAAGCTTGTGAAGACGAAAC---GCTCATGATA 4200
D 4170 TTACAGCCGATCAAGATAGGAAAGCGGACCTGTCAAGGTGCAAGCTAATATCATT 4229
Q 4201 CATGCTGAGAACCAATTTTCTAAGATGCCGAACCGAAGGCGACTTAAAGCTGCA 4260
D 4230 CATGCCGTAGGACCAAACTTCAACAAAGTTCCGAGGTTGAAGGTGACAAACAGTTGGCA 4289
Q 4261 GCTGCCATCAATGAGCATAAGGCTCATGTGTCAACGCTGAGGCGAATTCAAAAATATCAGTA 4320
D 4290 GAGGCTTATAGTCACTGCTGAAGATTTGCAACGATTAACAATGACATGACGATTT 4349
Q 4321 CCGCTACTGTCAACCGGACATCTAATCTGTGGCAAGATGAGATGATCATTTGCAAT 4380
D 4350 CCACTGTGTCCACCGGACATTTTCCGGGAACAAAGATGACATACCCATATTTGAAC 4409
Q 4381 CACCTGTCACTGTTTGCACACTACAGATGCGGATGTCAACATATTTCTTGATATA 4440
D 4410 CATTTGCTGACAGCTTTAGACACCACTGATGACAGATGTACCTCATATCTCAGGAGCAAG 4469
Q 4441 CAATGGAGACAGAAATATGAGGCAATTCACCGCAAAAGACGCTGAAATTTCTGGA- 4499
D 4470 AAATGGAAATGACTCTCAAGGAAGCAGTGGCTTAGGAGAAAGCAGTGGAGAGATATGC 4529
Q 4500 -----TGATGACAAGCCAGTAGACATTAAGCTTGGTCAAGGTTCAACCCAAAC 4545

QY 724 ACTAATAGATCAATATTCTCGGTTGGTTCAACATCTACAGAGAATAGTACTGTGA 783
DB 750 TCCAAAGATGTTCTATTCTCTGTTGGCTGACACATCTACAGAGAGAGGAACTTACTG 809
QY 784 CGTAGCTGGCAATCTTCCAAACGTTTCCACTTGAAGGAAAGTCTTACAGGTAGA 843
DB 810 AGAGAGCTGGGACCTGGCTGTATTTCACTTACGTGGCAAGCAAAATTCACATGTGCG 869
QY 844 TGTGGGACCAATTGCACTGTGGAAGGATACGTATCAAAAAAGAAAGATCAAGCCCAAGA 903
DB 870 TGTGAGACTATAGTTAGTTGGACGGGTACGTGTAAAGAAATAGCTATCAATGTCAGGC 929
QY 904 CTATACGCTAAAGTTGAGAACTTGGCGTCCAAATGATCGCGAGGTTCTTGAATTGC 963
DB 930 CTGTATGGGAAGCCTTCAAGGCTATGCTGTACATGCAACCGCGAGGATCTTGTGCTGC 989
QY 964 AAAGTCAAGATAGCTGCGCGGCGGAGAGGTTCTTTTGTGTGTGATCTATGTACA 1023
DB 990 AAAGTCAAGACATATTAACGGGGAGAGGGTCTTTTCCCGTGTGCAAGTATGTCCA 1049
QY 1024 GCCACACTTTGGCGATCAGATGACAGGAAATCTGTGGCACTGACGTATGTTGATGAGCA 1083
DB 1050 GCTACATGTTGTGACAAATGACTGTGCACTGTGCAACAGATGTGAGTGGGAGAGCGCG 1109
QY 1084 CAAAACTATTTGTTGGGCTCAACCAAGATGTGCTCAATGTGTAGGACGCAAAAGAA 1143
DB 1110 CAATACTGCTGTGTGGGCTCAACAGGTATAGTCTGCAACGGTCCGACCCAGAGAAAC 1169
QY 1144 ACTACCAATGCAAACTATCTATTAACAGTGTGCTGCGGCGGTTTCCAGGTGGCG 1203
DB 1170 ACCAATTCATGAAAAATTAACCTTTTGTCCGTAGTGGCCAGGCAATTTGCTAGGTGGCA 1229
QY 1204 CGTGAATCGTGTGCGACTTGGAGAGAGAAAGACTAGGGGGCGGGAGCGCACTCTT 1263
DB 1230 AAGGAATATTAAGAAATCAAGAAAGTAAAGGCCACTAGGACTACAGATAGACAGTTA 1289
QY 1264 ACTATGGGCTGTGCTGGGCTTTTCAAGACCCAGAAATTCACATCATCTACAAAGACT 1323
DB 1290 GTATAGGGGTGTTGTTGGGCTTTTGAAGGACAAAGTATACATCTATTTTAAACGCCCG 1349
QY 1324 GGTACGAAACAATTAAGAAAGTACGTGCGTCTTTGATCTGATTTGTGATTCAGCGCTT 1383
DB 1350 GATACCCAAACATCACTAAAGTAAACAGCATTTTCCACTATTCGTGCTCCAGAGATA 1409
QY 1384 ACAGCCACCGGCTCGATATGGGCTTCCGCGGTAGGCTCAAGCTGTCTGTAACCAACT 1443
DB 1410 GCGCATTAACATTTGGAGATCGGCTGAGAAACAAGAAATGTTAGAGAGACAC 1469
QY 1444 GTCAAAACCCGACCGGCTATTTACATGAGCCGATGTGAGCATCTGCGTGTACAGCAA 1503
DB 1470 AAGAGCCGTACCTCTCATTTACGCGCGAGAGACGTACAGAAAGCTAAGTGGCAGCCGAT 1529
QY 1504 GAAGCTGAAGATGTGGCTGACGCGGAAAGATTCAGAGAGCCCTGCGCACCTTGTCTCCT 1563
DB 1530 GAGGCTAAGAGGTGCTGTAAGCCGAGAGTGGCGCAGCTCTACACACTTTTGGCAGCT 1589
QY 1564 GAATATAAAAAGAGCGTATGAGGAGCAAGTACACCTCATTAAGCAAGAGCGAGACA 1623
DB 1590 GATGTATGAGGCCACTCTGGAAGCGATGTGACTGTATTAACAAGAGCTGGGGCC 1649
QY 1624 GGTAGCTGAGACACGACGACATCTAGGGTGAACAATTAACCAAGCGGAGAGAGAG 1683
DB 1650 GGCTCATGTGAGACACTCGTGGCTTGATTAAGGTACAGCTACGCTGGGAGAGACAG 1709
QY 1684 ATTGGGTCTTACGCTATATCTTTCAACCCAGCGGATTTGAATGAGAAAACTGGCGGT 1743
DB 1710 ATCGGCTCTTACGCTGTGCTTTCTCCGACGCTGTACTCAAGAGTGAATAATTAATCTTGC 1769
QY 1744 ATCCACCATTTGGGGAACAAGTACTGTATATGACTCAAAAGTAAAGGCGGAGATAC 1803
DB 1770 ATCCACCTCTCGCTGAACAAGTATGTATTAACACTCTGGCCGAAAGGCGGTTAT 1829
QY 1804 AAAGTGAAGCATACCAAGTAAAGTCAATGTATACAGAAAGGAGCGGCGTCCCTGTTCAA 1863

DB 1830 GCGGTGAACCATATACATGTTAAAGTATGTGTGCAGAGAGGACATGCAATACCGTCCAG 1889
QY 1864 GACTTCCAGGCAATTAAGTATGAGCGCTACGATGTTTCAACGAGAGGAGTTCTGTAAC 1923
DB 1890 GACTTTCACACTGTAGTGAAGTGCACCATGTGTATCAACGAACTGTAGTTCTGTAAAC 1949
QY 1924 AGATACCTGACCAACATCGCAATCAACGAGAGCGCTTAAACACTGACGAAGATATCTAT 1983
DB 1950 AGGTACTTGAACCATATTTGCTACACATGAGAGAGCGCTGAACACTGATGAAGATATTAC 2009
QY 1984 AAGACTGTAAAGACTCAGACACAGACTCAAGAAATACGCTTGTGATATTAAGCAGCAAG 2043
DB 2010 AAAAGTGTCAAGCCAGCGAGCAGACAGAGGGAATACCTGTATACGACATGCAAGAAACAG 2069
QY 2044 TGTGTTAAGCGAGAAAGACGAGGTCCCTTGTGCTTAACCGGTATCTGTATGATCCACA 2103
DB 2070 TGGCTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCAAGGCGAGCTGTGTGATCTTCC 2129
QY 2104 TTTCAAGAGTTTGGTACGAGAGTCTCAAGACAGACAGACAGACTCAAAAGTCCCA 2163
DB 2130 TTCTATATATTCGCTTACGAGAGTCTGAGAACAGACACACCGCTCTTACCAAGTACCA 2189
QY 2164 ACCATCGAGTCTATGAGAGTGCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGCTGTG 2223
DB 2190 ACCATAGGGGTATGAGCGTGTGCGCAGGATCAGGCAAGTCTGGCATCTTAAGAGCGCATC 2249
QY 2224 ACTTAAGAAATCTGTGTTGTGATGTGCGAAGAAAGGAAACCTGCCAGAAATCATCAGGAT 2283
DB 2250 ACCAAAAAATCTAGTGTGAGCGCCMAAGAAAGAAAGTGTGCAAGAAATTAATAGGAGAC 2309
QY 2284 GTAAGAGATGAGACGTATGATGTTGTCTGATAGACTGTGATCTGATCTGATCTTCTTAAT 2343
DB 2310 GTCAAGAAATGAAGAGGCTGAGCTCAATGCCAGAACTGTGACTAGTGTCTTGTGAAT 2369
QY 2344 GGGGTTAAGCAACCCCGTTAAACCTGTGTACATGTATGAGGCAATTTGCTGCTCATGAGG 2403
DB 2370 GAGTCAAAACACCCCGTATGAGACCTGTATATTGAACAACTTTTGTGCTGTATGACAGT 2429
QY 2404 ACGCTGTGCACTGATTTGCCATGCTCAAACTTAAGAAAGTGTGATGTGCGGGAGCCA 2463
DB 2430 ACTTTCAGACCGCTCATAGGCTATTAATTAAGACCTTAAGAAAGCAAGTCTCTCGGGAGTCC 2489
QY 2464 AAACAATGCGGCTCTTTTAAATGATATGCTGCTGAAGTAAATTTTAACTGATGACATATG 2523
DB 2490 AAACAGTCCGTTTTTTTAACTATATGTGCTGAAGTGTATTTTAACTGAGATTTGC 2549
QY 2524 ACTGAAGTATACATTAAGACATCTCTAGAGGTGACACAGACTGTAAACCGCATGCTC 2583
DB 2550 ACAAGTCTTCCACAAAGACATCTCTGCGGTTGCACTTAATCTGTGACTTGGGTGCTC 2609
QY 2584 TCCACGCTCTTCTTCAAGCAAGCGAATGAAGACGTTAACCCATGTGCTGATTAATATATC 2643
DB 2610 TCAACCTTGTTTTAAACGAAAAAATTAAGACAGCAAGATTCGAAGAGACTTAAGATTGTG 2669
QY 2644 ATGATATCAACAGGAGACCAAAAGCCGCAAAAGATATGATCTGATTTCTTAACCTGTTTGA 2703
DB 2670 ATTGACACTACCGGCACTACCAACCTAAGCAGAGCATCTCATTTCTCATCTGTTTGA 2729
QY 2704 GATAGGTGAAGACGCTACAGATTGACTACAAAAATCAAGAAATCATGACTGCGGCTGCA 2763
DB 2730 GGGTGGGTGAAGCAGTTGCAAAATGATTAACAAGGCAACAAATTAATGACGCGAGCTGCC 2789
QY 2764 TCGTAAGGACTTATAGCGGAAAGCGTTATATGCTGTCAAGTATCAAAAGTCAAGAGATCCA 2823
DB 2790 TCTCAAGGGGTGACCCGTAAAGGTGTATGCGTTGCTGTGAAGAGTGAATGAATAATCTT 2849
QY 2824 CTCTACTGAGACTTGTGAGACGTGAAGATGTTATCTTAACAGCAGAGAAAAAGCATT 2883
DB 2850 CTGTACGACACCTTACGAAATGTATGAAAGTCTTACTGACCGGACGAGAGCCGATC 2909
QY 2884 GTCTGAAGACGCTAGTGTGATCTCTGTGATTAAGCACTTACAGTAAATATCTCCGG 2943

D	2910	GTGTGGAAAA	CACTAGACCGCGGACCCCATGATTA	AAAAACATGACTGCCAGTACCTTCGGG	2969	
O	2944	GATTTACGCGCTT	CTATTGACGACTGGACGGCAACGACGCCATTTATGGAC	CGCGTT	3003	
D	2970	AATTTCACTGCGCA	CGATAGAGAGTGGCAACAGACATGATGTCATCATGAGCACATTC		3029	
O	3004	CTGTGTAAGCCG	CAGACAGCTGATGTGTTCCAGAA	TAAAGTGAACGTCTGCTGGCCGAA	3063	
D	3030	TTGGAGAACCGGA	CCCTTACCGACCTCTTCCAGAA	TAAAGGCAAAACGTGTGTGGGCCAAG	3089	
O	3064	GCTTTAGAGCAGT	CTTGGCCACCGGCCAAACTGTGGCTGACGAGACAGCAGTGGGAGACG		3123	
D	3090	GCTTTAGTGGCGG	TGCTCTTAAGACCGCTGGCATTA	GCATGACCACTGAACATTTGAAACACT	3149	
O	3124	TTGCACCCATTC	CAAGCATGACAGACGCGTACTACCTGAAATAGGACATGAATCTTCTTTGCG		3183	
D	3150	GTGGATTATTTT	TGAAAGGACAAAGAGCTCACTCAGAGAGATAGATTGMAACCACTATGC		3209	
O	3184	ACCAGGTTCTTT	GGAGTGAACCTGGACAGTGGGTTATTTTCCGCTCCTACCGTGCACCTT		3243	
D	3210	GTGAGGTTCTTT	GGACTCGATCTGGACCTCCGCTCATATTTCTGGACCCACTGTTCGGTTA		3269	
O	3244	ACTTAACAGGAT	TCAGCACTGGGATACTCGGCAAGGAAACATGTAATGGCTTAATAGA		3303	
D	3270	TCATTATGAAAT	TAATCACTGGGATTAACCTCCGTGCGCTTAACATGTAAGGCGCTGAATPAA		3329	
O	3304	GAGGTAGCAAG	AGATTGTCAACGGCGATATCCGTGCATCA	CAAAAAGCGGTGACACAGGC	3363	
D	3330	GAAGTGTCGCT	CAAGCTCTCTCGACGATACCCACAACTGCCTCCGGGACGTTTGCAC	CTGGA	3389	
O	3364	AGGGTAGCTGA	TATPAGGAATPACCATCAAGGACTATCTCTCCAA	CAATTAATGTGTT	3423	
D	3390	AGAGCTATGAC	ATGACACTGTGACACTGCGCAATTATGATCCGCGCATPAACTCTAGTA		3449	
O	3424	CCATTAAATGCC	CGGTTGCCCCCACTCGTTGATCTGTGACCA	CAAGGACAGGGTACACT	3483	
D	3450	CTGTAAACAGAA	AGCTGCTCTCATGCTTTAATGCTCCACCA	TAATGAACACCCACAGAGT	3509	
O	3484	GATCAACAGCG	GAATTCCTATCTTAAGATGAAAGGCAAA	CTGTGTGGATGATCGCGGATCT	3543	
D	3510	GACTTTCTT	CATTCTGTACGCAAAATGAAGGGCAGAACTGCTCTGTGTGTGGGAAAG		3569	
O	3544	ATCAGCATTC	CAGGAGAAAGTAGAGTCCATGSGTCCAT	TGCGCAATACCATCAGG	3603	
D	3570	TTGTCCGCTCC	AGGCAAAATGTTGACTGTGTTCAGACCGGCTGAGGCTTACCTT	CAGA	3629	
O	3604	TGTGATCTCGA	TTTGGGAATACCTAGCATGTCCGTAAT	TGACATTAATCTTTGTCAAT	3663	
D	3630	GCTCGGCTGGA	TTTAAAGGATCCCAAGTGATGTGCCAAAT	TGCAATTAATTTGTTAAT	3689	
O	3664	GTTTAGAGCCCC	CGGTACAGGAACATCACTACCAACAGTGCAGATCA	CGGTATCCACAC	3723	
D	3690	GTGAGAGACCC	ATTAATTAACATCACTATGACAGTGTGAAGCACTG	CCATTTAACTT	3749	
O	3724	AGCATGCTAA	CGTGAAGCTGTCCACCACCTGAA	CACTGGCGAAGCATGTGTGGCTATA	3783	
D	3750	AGCATGTGA	CCAAAGAAAGCTTTGCTGTGATCTGAA	TCCCGCGGAACCTGTGTAGACATA	3809	
O	3784	GGGTATGGGCT	TGCTGATCGGCAACGAGAA	TATCATCACTGCGGTGGCACGCTCA	TTT	3843
D	3810	GGTATATGTTA	CGCTGACAGGGCCAGCAAA	GCATCATTTGTGTCTAATAGCCCGGCA	AGTTC	3869
O	3844	AGGTTTACCCG	TGCTGTCAAGCTTAAGAA	CACTGCCGAAATATCTGAAGTTTCTTGGTG		3903
D	3870	AAAGTTTCCG	GGGTATGCAAAACCGAAATCTCACTTGAAGACGGAAGTTCTGT	TTGTGA		3929
O	3904	TTCTTGGCA	AGCAACGCGCAACCA	CACATGACAGGACAGACTCGGTGTA	AGTCTT	3963
D	3930	TTCTATGGGT	ATCGATCGAAGGCCGTATCGCAATCTTATCAAGCTTTATCA	CACTTG		3989
O	3964	GACAAACATCT	ATCAAAAGGTCAACAGAGTACGAGCAGGAGAGCTCCA	AGGATACAGATG		4023
D	3990	ACCAACATTTA	ACAGTTTCCAGACTCCACGAAACCGGATGTGACCTCATAT	CAATCATGTG		4049

QY	4024	ATCAGAGTGACATTAGCAAGAGCGCTGACCAAGCTATTCGTAAATGCTGCTAATAGCAAA	4083
Db	4050	GTGCGAGGGGATATTCGACAGGCCACCGAAGAGTATTAATAATGTGCTAAACAGCAAA	4109
QY	4084	GGTCAACCAAGGTTCCCGAGTGTGCGGTGCACTGTACCGAAATAATGCGCGCTGCTTTGAT	4143
Db	4110	GGAACAACCTGGCGGAGGGGTGTGGGAGCGCTGTATAAGAAAGTTCCCGGAAGCTTCGAT	4169
QY	4144	AGACAGCCAAATAGCTGTGCGGACGGCTAGACTTGTGAAGACGAAACC--GCTCATCATYA	4200
Db	4170	TTACAGCCGATCGAAGTAGAGAAAAAGCCGCACTGGTCAAAAGTGCAGCTAAACATATCATTT	4229
QY	4201	CATGTGTAGGACCCCAATTTTCTTAAGATGCGGGAACCGGAGGCGGACCTTAAGCTCGCA	4260
Db	4230	CATGCGCGTAGGACCAAACTTCACAAAGTTTCCGAGGTTAAGGTGCAAAAGTTGGCA	4289
QY	4261	GCTGCTTACATGAGCAATAGCGTCCATGTCAGACGCTGAGCGGATTACAAAATAATCATGTA	4320
Db	4290	GAGGTTTATGATGCATTCATCGTAAATTTGTCAACGATTAACAATTACAAAGTCAGTAGCATTT	4349
QY	4321	CCGCTACTGTCAACCGGATCTTATTTGTGTGGCAAGAGTGAAGTATCATCATTCGAT	4380
Db	4350	CCACTGTGTCCACCGGATCTTTTCCGGGAAACAAAGATGACATTAACCCATCATTTGAAC	4409
QY	4381	CACCGTTTCACTGCTTTTCGACACTAGAGGATGCCGATGTCACATATATTGCTTGATATAA	4440
Db	4410	CATTTGTGACAGCTTTTAGACACACTAGTAGCAGATTTAGCCATATATCTCAGGAGACAG	4469
QY	4441	CAATGGGAGACAGAGATATTCGAGGCCATTACCGCAAAAGAAAGCGTCGAATTTCTGGA-	4499
Db	4470	AAATGGGAAATGACTCTCAAGGAAGAGTGGCTTAGAGAGAAAGAGTGGAGAGATATATGC	4529
QY	4500	-----TGATGACAAAGCCAGTAGACATTGACTTTGGTCAGGGTTCACCCAAAC	4545
Db	4530	ATATCCAGCACTTTCAGTAGACAGAACTGATGACAAAGCTGTGTAGAGGAGTCATCCGAAG	4589
QY	4546	AGCTCTTTGGCAGGACAGACAGAGTTACTCCGTCAATGAGGGCAAGTTGTATTATACCTG	4605
Db	4550	AGTTCTTTGGCTGGAAGGAAGGCTTAGACACAAGCAATGGCAAACTTCTCATATTTGG	4649
QY	4606	GAAAGTACACGATTCCATCAGACCGGCCAAGAGACATTCGCAAAATCCATGACATGTGGCCC	4665
Db	4650	GAAAGGAGCAAAAGTTTCCACAGGCGGCCAAGATATAGCAAAATTAATGCAATGTGGCCC	4709
QY	4666	AACAAATCTAGGCTATATGAGCAGATTTGCTTGTAACATCTGGGGGAGAGCTATGTCCAGC	4725
Db	4710	GTTCGCAACGAGGCGCAATGAGCAGGATATGATATATCTTCGAGAAACATGACAGCTAGT	4769
QY	4726	ATCCGCTCCAATCCCAAGTAGAGAGTAAGAGGCGCTGCTCCACCTCACAACATTTCCA	4785
Db	4770	ATTAGGTGCAAAATCCCCCGTCGAAAGATCGGAAGCTCCACACACTAGACAGCTGCTCT	4829
QY	4786	TGCTGTGTAAATTAACGCTATGAGCGGCTGAGCGCGGTATACAGGTTGCGCTCTGCGAAGAAA	4845
Db	4830	TGCTTTGTGCATCCATGTCATGACTCCGAAAGAGTACAGCGCTTAAAGCTCACCTGCCA	4889
QY	4846	GAAACAGTTGCGCGATGCTCATATTCCTGTGTCGGAAGTACAGATACAGGCGGTGCGAG	4905
Db	4890	GAACAAAATTAATGTGTCTCATCTTTTCCATGTCGAAAGTATGAATATCACTGTGTGCGAG	4949
QY	4906	AAGCTTACAGTCAGCAAAACCAAGTCTGTATTTTTCAGGCGTCTGTACCAACGCGCTGTACCCC	4965
Db	4950	AAAGATCAAAATGCTCCAGCGCTATTTGTCTCACCGAAAGTGCCTGGTATATTATTCACA	5009
QY	4966	AGGAAGTA 4973	
Db	5010	AGGAAGTA 5017	

RESULT 4
AAD29139
ID AAD29139 standard; DNA; 12379 BP.

XX AAD29139;
AC
XX 07-MAY-2002 (first entry)
DT
XX
XX Human immunodeficiency virus type 1 (HIV-1) pol plasmid.
DE
XX
XX Human immunodeficiency virus; vaccine; HIV infection; immune response;
KM alpha-virus replicon; therapy; pol gene; de.
XX
XX Chimeric - Human immunodeficiency virus type 1.
OS
XX Chimeric - Unidentified.
PN
XX WC020203917-A2.
PD
XX 17-JAN-2002.
XX
XX 09-JUL-2001; 2001WO-US21701.
PF
XX
XX 07-JUL-2000; 2000US-216995P.
PR
XX
XX (ALPH-) ALPHAVAX INC.
PA (UTNC-) UNIV NORTH CAROLINA.
PI Olmsted R, Keith P, Dryga S, Caley I, Maughan M, Johnston R;
PI Davis N, Swenstrom R;
XX
XX WPI: 2002-171664/22.
DR
XX
XX Composition useful for treating or preventing HIV infections, comprises
PT two or more isolated nucleic acids encoding env, gag or pol gene
PT product of HIV or immunogenic fragment of the gene products -
XX
XX Example 5; Page 183-187; 201pp; English.
PS
XX
XX The invention relates to a composition comprising isolated nucleic acids
CC encoding env, gag or pol gene product of human immunodeficiency virus or
CC immunogenic fragment of the gene products. The gag gene product is
CC modified to inhibit formation of virus-like particles containing gag
CC gene product and their release from cells, and the pol gene product is
CC modified to inhibit reverse transcriptase activity. The invention also
CC relates to a method for producing an alpha-virus replicon particle used
CC in vaccines. The composition is useful for inducing an immune response
CC to human immunodeficiency virus (HIV) or for treating or preventing HIV
CC infection in a subject. The alpha-virus replicon particle is useful in a
CC vaccine. The composition is useful for administering a protein or peptide
CC to a subject. A composition comprising heparin affinity-purified alpha-
CC virus replicon particle is useful as a clinical trial material and as a
CC commercial product. The present sequence is HIV-1 pol plasmid.
CC
XX
SQ Sequence 12379 BP; 3562 A; 2920 C; 3084 G; 2813 T; 0 other;
Query Match 42.3%; Score 2113.2; DB 24; Length 12379;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3216; Conservative 0; Mismatches 1748; Indels 24; Gaps 3;

Db 270 CACAGTATCATGTATCTGCGATGAGATGTGCGGAGATCCGACAGATGTATTAAG 329
Qy 310 TATGCAAGAAAGCTTAAGAAAAGT-----GACATTACCAAGCAAGATATGACCTTAAG 363
Db 330 TATGCAACTAAGCTGAAGAAAGAAAGCTTAAGAAATTAATCTGAATTAAGATGAAGAA 389
Qy 364 GCGCAGACCTGTGTAAGTATGTCAACACAGACCGAGAGATCCATCTGTGTATG 423
Db 390 ATGAAGAGAGCTCCGCGCGCTCATAGAGACCTTGACCTGGAACTGAGACTATGTGCTC 449
Qy 424 CACACAGACCGCAGTGTAGTACTTTGGAAGTGTACAGTATACCAAGATGTATGCA 483
Db 450 CACAGACAGAGTGTGTGCTGCTACGAAAGGCAAGTCCGTTTACAGAGATGTATAGCG 509
Qy 484 GTCCATGACCGAATCAATCTACCAAGCGCTTAAGAGTTAGACAAATTTACTG 543
Db 510 GTTGACGAGCCGAAAGTCTTATCACCAAGCCAAATAGGAGTTAGATGCGCTACTG 569
Qy 544 ATAGCTTTGACAGACCCCTTTATGTACAAAACATGGCAGGTTCTTACCTACTTAC 603
Db 570 ATAGCTTTGACACACCCCTTTATGTATTAAGAACTTGCTGAGCATATCATATAC 629
Qy 604 AACAGAACTGGGCTGACGAGAGATTTGGAGACAGTAACTTGGCTTGTACTCA 663
Db 630 TCTACCAACTGGGCGACGAAACCGTGTAAAGGCTGTAAATAGAGCTATGACGCTCT 689
Qy 664 GATCTCAGAGAGACAGGCTTGGAAAACCTCATCTCTTGAAGAAAGAGCTTCAACT 723
Db 690 GACCTTATGAGAGCGGTACAGTATGAGGAGTGTCCATTTCTTGAAGAAAGTATTAACA 749
Qy 724 ACTATATAGATCATATTTCTGGTGTGTTCAACAATCTACAGAAATAGTATCATCTT 783
Db 750 TCCAACATGTCTATTTCTGTTGCTGACCATCTACAGAGAAAGAGGAACTTACTG 809
Qy 784 CGTAGCTGGCATCTTCCAAGTGTCCACTTGAAGAAAGTCTAACTTACAGGTAGA 843
Db 810 AGAGCTGGACACCTGCGTGTATTTCACTTACGTGGCAAGAAATTAACATGTGCG 869
Qy 844 TGTGCGACATTTGCACTGTGAGAGGTTGCTCATCAAAAAGATAGATAGCCCGAGA 903
Db 870 TGTAGACTATAGTATGTGTGAGCGGTAGCTGTAAAGAAATAGTATATGATGTCAGGC 929
Qy 904 CTATACGCTAAGTGTGAACCTTGCGCTCCACATGATGATGCGAGGTTCTTGTAGTGC 963
Db 930 CTGTATGGAAGCTTGAAGCTATGCTATGATGATGATGATGATGATGATGATGATG 989
Qy 964 AAAGTCAAGATGCTGCGCGGCGAGAGAGGTTTCTTTGCTGTGTATGATGATGATG 1023
Db 990 AAAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1049
Qy 1024 GCCACATTTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1083
Db 1050 GCTACATTTGTGTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1109
Qy 1084 CAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1143
Db 1110 CAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1169
Qy 1144 ACTAACCAATGACGAACATATCTATTAACAAGTGTGCGGAGGTTTCCAGGTGAGCG 1203
Db 1170 ACCAATGCAATGAATAATTAATCTTTGCGGTATGAGCCAGGATTTGTAAGTGGCA 1229
Qy 1204 CGTGAACATGTGCGGCTTGTGACGAGAGAAAGAACTAGGAGGTGCGGAGCGCATCTT 1263
Db 1230 AAGGAATATTAAGAGATCAAGAAAGTGAAGAGGCACTAGAGACTAGAGATGACAGTTA 1289
Qy 1264 ACTATGGCTGTGCTGCGGCTTCAAGACCGAGAAATCAATCATCTCAAGAGAGCT 1323
Db 1290 GTCATGGAGTGTGTGAGGCTTTTGAAGGACAAAGTATATTTATTAAGCGCGCG 1349
Qy 1324 GGTACGCAACATTAAGAAAGTACGTGCGTCTTTGACTATTTGTGATTCACGCGCTT 1383

Db 1350 GATACCCAAACCATCATCAAGTGAACAGCATTTCCACTATTCGTGCTGCCAGATA 1409
Qy 1384 ACCAGCCAGGGCTCGATATGCGCTTCCGCGTAGGCTCAAGCTGCTTGAAACCAACT 1443
Db 1410 GGCACTAACACATTGGAGATCGGGCTGAAGAACAGAAATCGAATAATGTTAAGAGCAC 1469
Qy 1444 GTCAAAACCCGACCGGCTATTACAAATGCGCATGTGAGACATCTGCGTGTTCAGCA 1503
Db 1470 AAGAGCCGTACCTCTCATTTACCGCCGAGAGCATCAAGAAAGCTAAGTGGCCAGCCAT 1529
Qy 1504 GAAGCTGAAGAGTGGCTGACCGGAAGAGATCAAGAAAGCCCTGCGACCTTCTCTCT 1563
Db 1530 GAGCTGAAGAGAGTGGCTGAAGCCGAGAGTGGCCGCGCTCTACCACTTTGCAAGCT 1589
Qy 1564 GAATAGAAAAAGAGACCGTAGAGGACAGAAATGAGACCTCATTTATGCAAGAGCAGAGA 1623
Db 1590 GATGTTGAGAGACCCACTCTGAGAGCCGATGTCGACTTGATTTACAGAGGCTGGGCCC 1649
Qy 1624 GGATGCGTGAAGACCAAGAGAGACATCAGGGTGAACAAGTTACCCAGCGCAAGAGAG 1683
Db 1650 GGCTCAGTGAAGACACTGTGGCTTGATTAAGGTTACCAAGCTACGATGGCGAGACAG 1709
Qy 1684 ATTGGCTTTAGGCTATACTTTACCCGCGGCTATTTGAATATGTAATAATCGCGCTGT 1743
Db 1710 ATCGGCTTTAGGCTGTGCTTTCTCGCAGGCTGATCTCAAGAGTGAATAATATCTTGC 1769
Qy 1744 ATCCACCATTTGGCGGAACAAGTACTGGTATGATGATCACTCAAAAGGTCAGGAGATAC 1803
Db 1770 ATCCACCTCTGCTGACACAGTCAATAGTATACACTCTGGCCGAAAAAGGCGTTAT 1829
Qy 1804 AAAGTCAGACATACCAAGGTAAGGTCAATTTGACAGAAAGGAGCGCGCTCTGTTCA 1863
Db 1830 GCCGTGAACCATACCATGTAAGTATGATGTCAGAGGACATGCAATACCCGTCCAG 1889
Qy 1864 GACTTCCAGGCAATTGATGAGAGCGCTACGATCGTTTTCACAGAGGAGTTGCTAAAC 1923
Db 1890 GACTTTCAAGCTCTGATGTAAGGTGCAACATTTGTATCAAGAAAGTGAATCTGTTAAC 1949
Qy 1924 AGATACCTGACCAATGSCAATCAACGAGAGAGCGCTTAACATGACGAAAGATCAT 1983
Db 1950 AGGTACCTGACCAATGTCACACATGAGAGAGCGCTGAACATGATGAATAATATATAC 2009
Qy 1984 AAGACTGAAGACTCAGAGACACAGACTCAGAATAGCTTCTGATATTTGACGACAGAAAG 2043
Db 2010 AAAATGTCAGAGCCAGAGAGACAGACGCGCAATACCTGTACGATCAGACAGAAACAG 2069
Qy 2044 TGTGTTAAGCAGAAAGACGACAGTCCCTTGTGCTTAACCGGTGATCTGTAGATCACCA 2103
Db 2070 TGCGTCAAGAAAGAACTAGTCACTGAGGCTACAGCGGACCTGGTGTATCTCCC 2129
Qy 2104 TTTCAAGATTGGCTAGAGAGTCTCAAGACAGACCGAGCACTTCAAAAGTCCCA 2163
Db 2130 TTCCATGAATTCGCTACGAGAGTCTGAACACAGACGCGCTCTTAACCAAGTACCA 2189
Qy 2164 ACCATCGAGTCTATGAGTGCAGAGTTCAAGTTAATCTGAATCATCAAAAGCCGTGTG 2223
Db 2190 ACCAATGAGGAGTATGAGCGTGCAGAGATCAGGCAAGTCTGGCATATTAAGAGCCGAGTC 2249
Qy 2224 ACTAAGAAAGATCTGTTGTGATGTCGAAAGAAAGAAACTGCGCAGAAATCATCAGGAT 2283
Db 2250 ACCAAAAAGATCTAGTGTGAGCGCCCAAGAAAGAAACTGTGCAGAAATTTAAGGGGAC 2309
Qy 2284 GTAAGAGAGATGAGACGATATGATGTTGCTGTAGAGCTGTGATTCAGTCTTCTTAAT 2343
Db 2310 GTCAAGAAATGAAGAGGCTGAGCGTCAATGCAAACTGTGACTGCTCTTTGAT 2369
Qy 2344 GGGGTTAAGCAACCCGTTAACAATCTGTACATTTGATGAGCAATTTGGCTGCATCAGAG 2403
Db 2370 GGATCAAAACACCCGTTAGAGACCTGTATATTTGAGCAAGCTTTTGTCTGTATCAGAGT 2429
Qy 2404 ACGCTCTGAGCACTGATTTGCAATGCTCAAACTTAAGAAAGTGTATTTGTGCGGGAACCA 2463
Db 2430 ACTTCAGAGCGCTCATAGCCATTTAAGAACTTAAGAAAGGCAAGTCTCTGCGGGATCCC 2489

Qy 2464 AAACATGCGGCTTCTTTAACAATGATGTCCTGAAGATACATTTTAACCATGACATATGC 2523
Db 2490 AAACATGCGGTTTTTTTAACTATGATGTCCTGAAGATGATTTTAACTACAGATTTTC 2549
Qy 2524 ACTGAAGTACCATTAAGACATCTCTAGAGAGTGCACACAGACTGTAACCGCATGCTC 2583
Db 2550 ACACAAGTCTTCCACAAAGACATCTCGCGCTTCACTTAATCTGTGACTTCGGTGTGTC 2609
Qy 2584 TCCAGGCTCTTACAGACAGGAATGAAAGACGTTAAACCATGTGCTGATTAATCATC 2643
Db 2610 TCACCTGTTTAAAGAAAAAATGAGAAAGACGAATTCGAAAGACATGAATTTGG 2669
Qy 2644 ATAGATACACAGGAGACCAAAAGCCGACAAAGATGATCTGATTTCACTGTTTCAGA 2703
Db 2670 ATTGACATACCGGAGATACCAAACTTAAGCAGAGACATCTCATCTCTGTTTCAGA 2729
Qy 2704 GGATGGGTGAACAGCTACAGATTGACTCAAAATTCAGAAATCAGAAATCAGTCCGCTGA 2763
Db 2730 GGGTGGTGAAGCAGTTGCAAAATGATTAACAAAGGCAAGAAATATGACGCGAGCTGCC 2789
Qy 2764 TCGCAAGACTTACCGCGGAAAGCGTTATGCTGTACAGTACAAAGTCAAGAGATCCA 2823
Db 2790 TCTCAAGGAGTGAACCGTAAAGGTGTATGCGCTTCCGTACAGGTGAATGAATCCCT 2849
Qy 2824 CTCTACGACAGACTTCTGAGACGTTGACGATGTAACGTTACTTACAGCAGAAAAAGCATT 2883
Db 2850 CTGTACGACCCACTTCAGAAATGTAAGTGTCTTCTGACCCGACGAGAGACCGCATC 2909
Qy 2884 GTCTGAAGAGCGTACGCTGATGATCCCTGATTAAGACATTAAGACTTAATATCCCGG 2943
Db 2910 GTGTGAAGAAACATGACCGGGGAGCCATGATTAAGAAACATGACCTGCCAAGTACCTGG 2969
Qy 2944 GATTTACAGGCTTCAATTGAGACATGAGACCGGCAACAGAGCCATTAATGSCAGCTT 3003
Db 2970 AATTACCTGACCATGAGAGAGTGAAGAGCAGAGCATGATGATGATGAGGACATC 3029
Qy 3004 CTGTAAAGCCGACAGACGCTGATGTTTCCAGATTAAGTGAAGTGTGCTGCGGCGAAG 3063
Db 3030 TTGAGAAACCGGACCTTACCGACGCTTCCAGAAATTAAGCAAGTGTGTTGGCGAAG 3089
Qy 3064 GCTTGAAGCAGATCTTGCCACGCGCAACATTTGCTGACGAGACAGCAGTGGGAGAG 3123
Db 3090 GCTTTAGTGGCGGTCTGAAGACCGCTGACATGACATGACCACTGAACAAATGAAACACT 3149
Qy 3124 TTGACCAATTCAGATGACAGACGCTACCTGAATGSCACTGAACTTTCTTTTGC 3183
Db 3150 GTGATTTATTTGAACCGACAAAGCTCACTCAGCAGAGATGATTTGAACCAACTATGC 3209
Qy 3184 ACCAGTCTTTGAGTGAAGCTTGAACAGTGGGTTATTTCCGCTCTACCGTGCACCT 3243
Db 3210 GTGAGTCTTTGAGCTGATGATGATGATCCGCTTATTTTCTGACCCACTGTTCCTTTA 3269
Qy 3244 ACTTACAGGATACAGACTGGGATTAATCTGCGCAGGAGAAACATGATATGGCTTAATGA 3303
Db 3270 TCCATTAAGAAATTAATCACTGGGATTAATCTCCGCTGCTTAACATGTAAGGCGTGAATAA 3329
Qy 3304 GAGGTGAAGAGAGTGTCAAGCGATATCTGTGATCACAAAAGCGTTGACACAGGC 3363
Db 3330 GAAGTGTTCGCTAGCTCTCTGACAGTATCCCAAACTGCGCTCGGCGAGTTCCACTGGA 3389
Qy 3364 AGGATGCTGATTAATGAAGATTAATCACTCAAGACTACTCTCCAACAATTAATGTGTT 3423
Db 3390 AGAGCTATGACATGAACATGTAACCTGTAACCTGCGCAATTAATGATCCGCGCAATTAACCTAGTA 3449
Qy 3424 CCATTAATTCGCGGTTGCTCCCACTGTTGATCGTTGACCAAAAGGACAGGGTCAACT 3483
Db 3450 CCTGTAACAGAAAGCTGCTCATATGCTTTAATGCTCCACAAATTAAGAACACCAAGAGT 3509
Qy 3484 GATCAGAGGATTTCTATCTTAAGTGAAGGCAAAATCTGTGTTGTGATTCGGCATCTT 3543
Db 3510 GACTTTTCTTCAATTCGTACGCAAAATTAAGAGGCAAGATGCTGTGCTGGGAGAAAG 3569

```

Qy 3544 ATCAGCATTCGAGGGAAGAAAGTAGAGTCATGGGTCCATTTGCCACTAATACATCAGG 3603
Db 3570 TTGTCGTCGCCAGGCAAAATGTTGATGCTGTTGTCAAGCCGGCTGAGGCTACCTTCAGA 3629
Qy 3604 TGTGATCTCGATTTGGGAATACCTAGCCATGTCCGTAATATGACATTAATCTTTGTCAAT 3663
Db 3630 GCTCGGCTGGATTTAAGGATCCGAGTGATGTGCCAAATATGACATTAATTTTGTTAAT 3689
Qy 3664 GTTAGGACCCCGCTCAGGAACCATCACTACCAAGTCGAGATCAAGCATCAACAC 3723
Db 3690 GTGAGAACCCCATTAATTAATCACTATCAAGCATGTGTAAGACCATGCTTAATAGCTT 3749
Qy 3724 AGCATGCTAACGTGTAAAGCTGTCCACCCTGAACACTGCGGAACATGTGTGCTATA 3783
Db 3750 AGCATGTTGACCAAGAAAGCTGTGTGATCTGAATCCCGCGGAACCTGTGTGAGCTA 3809
Qy 3784 GGGTATGGGCTTGTCTGATCGGCAACCGAATATCATCTGCGGTGGACGCTCATTT 3843
Db 3810 GGTATGGTTACGCTGACAGGGCCAGCGAAGCATCATTTGCTATATGCGCGCAGTTTC 3869
Qy 3844 AGTTTACCCGCTGTCTGCAAGCTTAAGAACCTGCCGAATATCTGAGGTTCTGTGCTG 3903
Db 3870 AAGTTTTCGCGGATGTAAGAACCCGAATCTCACTTGAAGACGGAAGTTCTGTGTTGTA 3929
Qy 3904 TTCTTCGCAAGGACCAACGGCAACCAACATGACAGACAGACTCGGTGTAGTGCTT 3963
Db 3930 TTCAATGGGTACGATCGCAAGCCCGTACGCAATCTTACAAAGCTTTATCAACCTTG 3989
Qy 3964 GACAACTATCTCAAGGCTCAACAGGTAAGAGGAGGAGAGCTCCAGCGTACAGAGTG 4023
Db 3990 ACCAACTATTAACAGGTTCCAGACTCCACGAAGCCGATGTGCACTCATATCATGTG 4049
Qy 4024 ATCAGAGGTGACATTAGCAAGAGGCTGACCAAGCTATCGTTAATGCTGTAATAGCAAA 4083
Db 4050 GTGCGAGGGGATATTCACCGCCACCGAAGAGTATTAATTAATGCTGCTACACGAAA 4109
Qy 4084 GGTCAACACGAGTTCCGAGTGTGCGGTGCTGTAACCGAAATGCGCGGCTGCTTTGAT 4143
Db 4110 GGACAACTGTGCGGAGGGGTGTGCGAGCGCTGTATTAAGAAATTCGCGGAAGAGTTGAT 4169
Qy 4144 AGACAGCCATAGCTGTGCGGACGCTAGACTTTGTAAGACGAAAC--GCTCATCTATA 4200
Db 4170 TTACAGCCGATCGAAGTAGAAGAAAGCGGACTGTCTAAAGGTGACGCTAAACATATCAT 4229
Qy 4201 CATGCTGTAGGACCCCAATTTTCTAAGATGCGGGAACCGGAGGCGACCTTAACCTGCA 4260
Db 4230 CATGCCGATGAGCAAACTTAACAAAGTTTCGAGGTTGAAGTGAACAAACAGTTGCA 4289
Qy 4261 GCTGCTTACATGAGCATAGCTGCATGCTCAACGCTGAGCGGATTAACAAAATATCACTA 4320
Db 4290 GAGGCTTATGATGTCATCGTCAAGTTGTCAACGTAACATTAACAGTCAGTAGCGATT 4349
Qy 4321 CCGCTACTGTCAACCGGATTTATTTGTGTGCGAAGATGAGTGAATCATTTGAT 4380
Db 4350 CCACGTGTGTCACCGGATTTTTCGCGAACAAGATCAGCTAACCCATCATTTGAAC 4409
Qy 4381 CACCTGTTCACTGCTTTCGACACTACGAGATGCCGATGACCATATATTTGTTGATATA 4440
Db 4410 CATTTGCTGAGAGCTTTAAGACACACTGATCAGAGCTGTGAGGGTGCATCCGAG 4469
Qy 4441 CAATGGAGAGACAGAGTATTCAGAGGCCATTCACCGCAAGAAAGCGTGAATTTCTGA- 4499
Db 4470 AAATGGGAATGATCTCTCAAGGAAGAGTGTGAGAGAGAGAGAGATGATG 4529
Qy 4500 -----TGATGACAGCCAGTAGACATTTGCTGTGAGGGTCCACCCAAAC 4545
Db 4530 ATATCCAGCAGCTTTCAGTATACAGAACTGATCAGAGCTGTGAGGGTGCATCCGAG 4589
Qy 4546 AGCTTTTGGAGGACGACGAGTTATCTCCCTCAATAGGCGCAAGTTGATCTACTG 4605
Db 4590 AGTTCTTTGCTGGAAGGAGGGCTACAGCAAGCGATGCGCAAACTTTCTCATATTTTG 4649
Qy 4606 GAAGGTACAGATTCATCAAGACGCGCAAGAGCATTTGCCAAATTCATGCAATGTGCGC 4665

```

```

Db 4650 GAAGGACCAAGTTTCAACAGCGCGCCAGAGATATACAGAAATTAATGCAATGTGCCCC 4709
Qy 4666 AACAATCTAGGCTAATGACGATTTGCTGTGACATCTCGGGGAGAGATATGCGAC 4725
Db 4710 GTTGCAAGGAGGCAATGAGCAGATGATGATATCTCGGAGAAAGCATAGACAGT 4769
Qy 4726 ATCCGCTCCAAATGCCAGTAGAGAGTCAAGAGCGTGTGCTCCACTCACACTTCCA 4785
Db 4770 ATTAGTTCGAATATCCCGTTCGAAGAGTCCGAAGCTCTCTACACCTAGACGCTGCT 4829
Qy 4786 TGCCTGTATTAATACGTTATGACGCGTGAAGCGGTATACAGTTGCTGCGAAGAA 4845
Db 4830 TGCTTGTGACATCCATGCGCATGACTCCAGAAAGATACAGGCGCTTAAAGCTCAAGTCA 4889
Qy 4846 GAACAGTCCCGGATGCTCATCATCTCTGTGCGGAAGTACAGGATCACAGGCGTGAG 4905
Db 4890 GAACAAATTAATCTGTGCTCATCTTCTTCATGCGAAGTATGAATCACTGGTGTGAG 4949
Qy 4906 AAGCTACAGTGCAGCAACAGTCTGTTTTCAGGCGCTGTACCAAGGCTGTACACCC 4965
Db 4950 AAGATCCAAATGCTCCAGGCTATATTTGTTCTCACCGAAAGTGGCTGTATATCAATCA 5009
Qy 4966 AGGAAGTA 4973
Db 5010 AGGAAGTA 5017

```

```

RESULT 5
AAD29141
ID AAD29141 standard; DNA; 13584 BP.
XX
AC AAD29141;
XX
DT 07-MAY-2002 (first entry)
XX
DE Venezuelan equine encephalitis virus (VEE) RNA replicon plasmid.
XX
KW Venezuelan equine encephalitis virus; VEE; vaccine; HIV infection;
KW human immunodeficiency virus; alpha-virus replicon; immune response;
KW therapy; ds.
XX
OS Chimeric - Venezuelan equine encephalitis virus.
XX
PN WO200203917-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US21701.
XX
PR 07-JUL-2000; 2000US-216995P.
XX
PA (ALPH-) ALPHAVAX INC.
PA (VUN-) UNIV NORTH CAROLINA.
XX
PI Olmsted R, Keith P, Dryga S, Caley I, Maughan M, Johnston R;
PI Davis N, Swanstrom R;
XX
DR WPI; 2002-171664/22.
XX
PT Composition useful for treating or preventing HIV infections, comprises
PT two or more isolated nucleic acids encoding env, gag or pol gene
PT product of HIV or immunogenic fragment of the gene products
XX
PS Example 5; Page 190-195; 201pp; English.
XX
CC The invention relates to a composition comprising isolated nucleic acids
CC encoding env, gag or pol gene product of human immunodeficiency virus or
CC immunogenic fragment of the gene products. The gag gene product is
CC modified to inhibit formation of virus-like particles containing gag
CC gene product and their release from cells, and the pol gene product is
CC modified to inhibit reverse transcriptase activity. The invention also

```

relates to a method for producing an alpha-virus replicon particle used in vaccines. The composition is useful for inducing an immune response to human immunodeficiency virus (HIV) or for treating or preventing HIV infection in a subject. The alpha-virus replicon particle is useful in a vaccine. The composition is useful for administering a protein or peptide to a subject. A composition comprising hepatitis affinity-purified alpha-virus replicon particle is useful as a clinical trial material and as a commercial product. The present sequence is Venezuelan equine encephalitis virus (VEE) RNA replicon plasmid used in the exemplification of the invention.

Sequence 13584 BP; 3942 A; 3114 C; 3392 G; 3136 T; 0 other;

Query Match 42.3%; Score 2113.2; DB 24; Length 13584;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3216; Conservative 0; Mismatches 1748; Indels 24; Gaps 3;

```
QY 10 ACTAATCATTCATATATGAAAGAAATTCACGTTGACTTAGATGTCGACGCCGTATGTC 69
DB 30 AATTACCTAACCCAAATGAGAAAGTTGACGTTGACATCGAGAGACAGCCCATTCCTC 89
QY 70 AAGGTGTAACAGCGGACGTTTCCACATTTGATGTCGAGCAAGGACAGTCACTACAT 129
DB 90 AAGCTTTCAGCGGAGCTTCCCGCAGTTTGAGTAGAAGCCAGACAGTCACTATAT 149
QY 130 GACCATGCAATGACGAGACGCTTTTCGATGTCGCAACAAAGCTCATTGAGAGCGAAGTC 189
DB 150 GACCATGCTAATGCCAGAGCGTTTTCGATCTGCTTCGCTTCAAACTGATGAAACGAGGTG 209
QY 190 GACCGGAGCAAGTTATCTTGACATTTGAGAGTGGCGCCGCTGACATGCACTTCCAT 249
DB 210 GACCATTCGACACCATCTTGAACATTTGAGAGTGGCGCCGCGCAGATGATTTCTAAG 269
QY 250 CACCGCTATCATTTGATCTGCTGATGATTAAGCGGTGAAGACCCGAGACAGTCAAAAGG 309
DB 270 CACAAGTATCATTTGATCTGCTGATGATGATGAGAGATGCGGAGACGATGATTAAG 329
QY 310 TATGAGAAAGACTTAAGAAAAAGT-----GACATTACGACAAAGACATAGGCTCTAAG 363
DB 330 TATGGAATTAAGCTGAAGAAAACTGTAAGAAATATCATGATAGAAATGAGCAAGAAA 389
QY 364 GCGGACAGCTGCTGGAAGTCAATGCAACCAAGCGACAGACATCTCTCTGTGTATG 423
DB 390 ATGAAGAGAGCTGCGCGCGTATGAGCGACCTGACCTGAGAACTGAGATATATGCGTC 449
QY 424 CACACAGCGCCACGTGATGATCTTTGGAAGTGAAGATATACCAAGATGTATGCCA 483
DB 450 CACGACGAGAGTCTGTCTGCTTACCAAGGCGCAAGTCTGTCTTACAGAGATGATATACCG 509
QY 484 GTCCATGACCGGACATCAATCTACCAACGAGCGCTTAAAGGATGAGCAATTTACTGG 543
DB 510 GTTGAAGGACCGACAGTCTCTATACCAAGGCAATTAAGGAGTTAGAGTGCCTTACTGG 569
QY 544 ATAGGCTTTGACAGCAACCCCTTTTATGTAACAAAAATGCGAGTTCCTTACTTATAC 603
DB 570 ATAGGCTTTGACACCAACCCCTTTTATGTTAAGACTTGGCTGAGCATATCCATCATAC 629
QY 604 AACACGAATCGGCTGAGAGAGATATTGGAAGACGTAACTTGGCTCGGTGATCTCA 663
DB 630 TCTACCAATCGGCGGAGGAAACCGTGTAAACGCTGTAACTATGCGCTTATGACGCTCT 689
QY 664 GATCTTTCAGAGAGAGGCTTGAAGAACTCTCAATCTTTAGGAAAGAGGCTCAACT 723
DB 690 GACGTTATGAGCGGCTACGTAAGAGGATGTCATTTTAAAGAAAGTATTTGAAACCA 749
QY 724 ACTAATAAGATCATATTTCTGGTTGGTTCAACAACTTACACAGAGATAGATCACTGTTA 783
DB 750 TCCAACATATGTTCTATCTCTGTTGGCTGACCATCTACACAGAGAGGGAATTAATCG 809
QY 784 CGTAGCTGGCATCTTCCAAAGTGTTCACATTTGAAAGAAAGTCTTACTTCCAGAGTGA 843
DB 810 AGGAGCTGGCACCTGCGCTGTGTATTTACATTAACGTGCAAGCAAAATTAACACATGTGG 869
```

```
QY 844 TGTGGACCAATTCAGCTGTGAAAGGTACGTCAATCAAAAAAGATACATGACCCAGGA 903
DB 870 TGTGAGACTATATAGTTAGTTGGACGGGACGTGCTTAAAGAAATAGCTATGTCACAGC 929
QY 904 CTATACGCTAAAGTTGAGAACTTGGCGTTCACATGCTCCGAGAGGTTTCTTGAGTTGC 963
DB 930 CTGATGGAAGGCTTCAGGCTATGCTCTACATGACACCCGAGAGATCTTGTGCTGC 989
QY 964 AAAGCACAGATPACGCTGCGGCGAGAGGGTTTCTTTGGCTGCTGATACGATATACCA 1023
DB 990 AAAGTACAGACACATTTCAACGGGAGAGGGTCTTTTCCGCTGTCACGATATGTCCA 1049
QY 1024 GCCACACTTTCGATCAGATGACAGGAGTTCGTGCACTGACGTTAGTGTGATACGCA 1083
DB 1050 GCTACATTTGTGTGACCAATATGCTGGCATATCTGGCAACAGATGATAGTGGGACGACGG 1109
QY 1084 CAAAACTATTTGTTGGGCTCAACCAAGAAATTTGCTCAATGTGAGACGCAAGAAAT 1143
DB 1110 CAAAAACTGTGTTGGGCTCAACCAAGGATATGCTCAACGCTGCAACCGAGAGAAAC 1169
QY 1144 ACTAACACAAATGCAAACTATCTATTACAGTGTGCGCCAGGCGTTTCCAGTGGCG 1203
DB 1170 ACCAATACCAATGAAATAATTAATCTTTTGCCTGATGSCCAGGCAATTTGCTAGTGGCA 1229
QY 1204 CGTGAACATCGTGCCGACTTGAACGACGAGAAAGAACTAGAGGCTCGGAGGCACTCTT 1263
DB 1230 AAGGAATTAAGAAAGATCAAGAAAGTAAAGGCCATAGACATACGAGATAGACATTA 1289
QY 1264 ACTATGGCTGCTGCTGGGCTTTCAAGACCAAGAAATCAATTCATCTACAAAGCT 1323
DB 1290 GTCAATGGGCTGTTGTTGGCTTTTGAAGAGCAAAAGATTAATTAATTAAGCGCCG 1349
QY 1324 GGTACGCAAAATTAAGAAAGTACCTGCGCTTTGCTCATTTGATATTCAGGCTT 1383
DB 1350 GATACCCCAACATCATCAAGAGTAAACGCAATTTTCACTCATTTGCTGCTGCCAGATA 1409
QY 1384 ACCAGCCAGGCTCGATATGAGGCTTCCGCGTACGCTCAAGCTGCTCTTGAACCAACT 1443
DB 1410 GGCAGTAAACATTTGAGATCGGCTGAGAACCAAGAAATCAGAGAAATTTTGAAGAGCAC 1469
QY 1444 GTCAAAACCCGACCGGCTATTATTAACAATGCGCATGTGAGCATCTGCGCTTACAGCA 1503
DB 1470 AAGGAGCGGTACCTCTCATTTACCGCGAGAGACGTACAAGAGTAAAGTGCACCGAT 1529
QY 1504 GAACTGAAGAAAGTGTGCTGACGCGGAGAGATCAGAGAAAGCCCTGCCACCTTCTCCT 1563
DB 1530 GAGGCTAAGAGAGTGTGAGAGCCGAGAGAGTGTGCGCGCATTAACCTTTTGGCAGCT 1589
QY 1564 GAAATTAAGAAAAAGACCGTGAAGGACAGAGTACCTCATTTGCAAGAGGACGAGACA 1623
DB 1590 GATGTTAGAGGCCCATCTTGTGAAGCCGATGTCACTTGATGTTTAAAGAGCTGTGGGCC 1649
QY 1624 GGTAGCGTGAAGACCAACGAGACACATCAAGGTGACAACTTACCAGGAGAGAGAG 1683
DB 1650 GGTCAAGTGAAGACACCTCGTGGCTGTGATTAAGGTATCACACTACGATGGGAGAGACAAG 1709
QY 1684 ATTTGGCTTACGCTATATCTTTCACCCGAGCGGTATTAATGTAAGAAAACTGCGCTGT 1743
DB 1710 ATCGGCTCTTACGCTGTGCTTCTTCCGAGGCTGTACTCAAGAGTGAAGAAATTTATCTTGC 1769
QY 1744 ATCCACCATTGCGGAGAAACAGATCTGGTATATGCTCAACAAAGTGAAGGACGAGATAC 1803
DB 1770 ATCCACCTCTCGCTGAACCAAGTCAATGATGATTAACACCTTGGCCGAAAAGGGGCTTAT 1829
QY 1804 AAAGTCAGACCATACCAAGTAAAGTCAATTTGACAGAGAGGAGCGCGCTCTGTTCA 1863
DB 1830 GCCGAGAACCATATCAATGTAAGTGTGTGCAAGGAGACATGCAATACCGCTCAG 1889
QY 1864 GACTTTCAGGATTTGATGAGAGCGCTACATGCTTTTCAACGAGAGGATTTGTAAC 1923
DB 1890 GACTTTCAAGCTCTGAGTGAAGAGTGCACATTTGTATCAACGAGACGATTTGTAAC 1949
QY 1924 AGATACGTCACCATGTGCAATCAACGAGAGCGCTAAACATGACGAAGATATAT 1983
```


1950 AGGACTGACCACTATTTGCCACACATGAGAGAGCGGTGAACATGATGAAGATATTAC 2009
1984 AAGACTGAAGACTCAGAGACACAGACTCAGATTAAGTCTTGATATTGACGACGAAG 2043
2010 AAAAAGCTGCAAGCCAGAGAGACAGAGCGGCAATACCTGTACGATCGACAGGAACAG 2069
2044 TGTGTTAAGGAGAAAGACGAGGTCCTTGTGCTTAACCGGTGATCTGTGATATCCACA 2103
2070 TGCCTCAAGAAAGAACTAGTCACTGAGGCTCAGAGGCTCAGAGGAGCTGTGATCTCC 2129
2104 TTTACAGAGTTTGGCTACGAGAGTCTCAAGACAGACGAGGAGGAGCTCAAGATCCCA 2163
2130 TTGCATGAATTCGCTACGAGAGTCTGAGAACAGACAGACGCGCTCTTACCAAGTACA 2189
2164 ACCATCGAGTCTATGAGAGTCCAGGTTCAAGTAAATGTGAATCATCAAAAGCGCTGTG 2223
2190 ACCATAGGGGTGTATGGGTGCTCCAGAGTCAAGCAAGTCTGGCATCATTAAGGCGCAGTC 2249
2224 ACTAAGAAAGATCTGTTGTGAGTGCAGAAAGGAAAACTGGGAGAAATCATAGGGAT 2283
2250 ACCAAAAAGATCTAGTGTGAGCGCCAGAAAGAAATCTGCAAGAAATTAAGGAGAC 2309
2284 GTAAGAGATGAGACGTATGATGATGCTGCTAGAGCTGTGATTCAGTCTTCTTAAT 2343
2310 GTCAAGAAATGAAGAGGCTGAGAGTCAATGCCAGAACTGTGACTGAGTCTCTTAAT 2369
2344 GGGGTTAAGACCCCGTTAACACTCTGTACATTGATGAGGATTTCCTGCTCAGTACAGG 2403
2370 GGATGCAAAACCCCGTAGAGACCCCTGTATTGTGCGAAGCTTTGTGCTGTACAGAGT 2429
2404 ACGTGTGCTGCACTGATTCGATGCTGCAACCTAAGAAAGTGTATGTGGGGGAGCCCA 2463
2430 ACTTCAGAGGCTCATAGCTATTAAGCTTAAGAAAGGAGAGTCTGTGGGGATCC 2489
2464 AACAATGCGGCTCTTTAATCATGATGCTCTGAAAGTACATTTTAACCATGACATATGC 2523
2490 AAAAGAGCGGTTTTTTAATGATGTCCTGAAGTGCATTTTAACAGAGATTTGC 2549
2524 ACTGAAGTATACCAATAAAGATCTTAGAGGTGCACACAGACTGTAAACCGCATGCTC 2583
2550 ACAAGAGTCTTCCAAAGAGATCTCTGCGGTGCATTAATCTGTGACTTCGCTCCTC 2609
2584 TCCAGGCTCTCTGACAGAGAAAGGAAAGGTTAACCATGTCGATTAATATCATC 2643
2610 TCAACTTGTTTTACGAAAAAATGAAAGAGAGATCCGAAGACATGAAGTGTG 2669
2644 ATAGATACCAAGGAGACCAAGAGCCGCAAAAGTATGATTTCTAACCCTGTTTTAGA 2703
2670 ATTGACACTACCGGAGTACCAAACTTAAGAGAGAGATCTCATTTCTCACTTGTTCAGA 2729
2704 GGATGGGTGAACAGCTACAGATGATCAAAAAATCAAGAAATCATGATCGGCTGCA 2763
2730 GGGGTGGGTGAAGAGTGGCAATGATTAACAAAGCAAGAAATATATACCGCAGCTGCC 2789
2764 TCCGAAAGACTTAGCGGGAAGGCTTATGCTGTCAAGTCAAAAGTCAAGAGAAATCCA 2823
2790 TCTCAAGGCTGACCCGTAAGGTGTGTATGCTGCTGTAACAGTGAATGAATCTT 2849
2824 CTCTACTGACAGACTTTGACAGACGTGAACGTGTTACTTACACGACAGAAAAAGCATT 2883
2850 CTGTACGACCAACCTCAGAAATGTGAACGTCTACAGACCCGAGAGGAGCCGATC 2909
2884 GTCTGGAAGAGCTAGCTGTGTATCTCTGATTAAGACATTTACAGCTAAATATCTCCGG 2943
2910 GTGTGGAAGAACTAGCCGAGACCATGATTAAGAAACATGACTGACCAAGTACCTTGG 2969
2944 GATTTCACGGCTTCAATGGAAGACTGAGAGCGGCAACGAGCCATTAATGAGCAGCTT 3003
2970 AATTTCATGACAGATGAGAGTGTGCAAGAGCATGATCCATATAGAGGACATC 3029
3004 CTTGATTAAGCCGACAGACTGATGTGTTCCAGATTAAGTGAAGCTGTGCTGGCGAAG 3063

3030 TTGAGAGACCGGACCTTACGACGTCTCCAGATTAAGCAAAAGTGTGTGGCCCAAG 3089
3064 GCTTTAAGCAGCTCTTGCCACAGCCCAAGATTGTGCTGACGAGACAGTGGGAGACG 3123
3090 GCTTTAGTGCCTGGCTGAAGACCGCTGAGCATGACATGACCACTGAACATGGAACCT 3149
3124 TTGCAACCATTTCAAGCATGACAGAGCTACTCACTGAAATGAGCATGAACTTCTTTTGC 3183
3150 GTGATTAATTTTGAAGGACCAAAAGCTCATCTAGCAGAGATGATTAAGAACAACTATGC 3209
3184 ACCAGGTTCTTTGAGTAGACCTGAGACATGAGGTTATTTTCCTCTTACCGTGCATT 3243
3210 GTGAGGTTCTTTGAGCTGATCTGAGCTCCGAGTCTATTTTCTGACCCACCTGTCCGTTA 3269
3244 ACTTACAGGAGTACGACTGGGATTAATCCGACGGGAAGAACATGTATGGGCTTAATGA 3303
3270 TCCATTGGAATATATCATGAGGATCTCCCGTGCCTTAACATGTAAGGGCTAATPAA 3329
3304 GAGGTACAAAGAGTGTGACCGGCGATATCCGTGCATCACAAAAGCGTTGACACAGCG 3363
3330 GAAGTGTCCGTGACGCTCTCCGAGGTACCCACAACTGCTCGGCGAGTTGCCACTGGA 3389
3364 AGGTAAGCTGATATAGGAATATATACATGAGACTTCTTCAACATTAATGTGGTT 3423
3390 AAGTCTATGACATGAACATGATACCTGCACTGCAATATATATCCGCAATPAACTAGTA 3449
3424 CCATTAATGCGCGTGGCCCACTCGTGAATGCTTGAACCAAGGACAGGGATCAACT 3483
3450 CCGTAAACAGAAACGCTCATGCTTTAGTCTTCCACATTAATGAACCCACAGAGT 3509
3484 GATCAGACCGGATTCCTATCTAAGATGAAGGCAAACTGTGTGTATGCGGATCCT 3543
3510 GACTTTCTTTCATTCGTGACGAATTAAGGGAGAACTGTCTGTGTGTGGGAAAAAG 3569
3544 ATCAGCATTCAGGAGAAAGTATAGATCCATGAGTCAATGCTTCCATTAATCATGACG 3603
3570 TTGTCCGTCCAGGCAAAAGTGTGATGCTGTGTCAGACCGGCTGAGGCTTACCTTACA 3629
3604 TGTGATCTCATTTGGGAATATCTAGCATGTGCTGAATATGACATTAATCTTGTCAAT 3663
3630 GCTGGCTGATTTAAGCATCCAGTGAATGTGCCAAATATGACATTAATTTGTATAT 3689
3664 GTTAGACCCCGTACAGAAACATCATCAACACAGTGCAGAGTACAGCTATCACACAC 3723
3690 GTGAGAACCCATTAATTAATCATCATCATCATGAGAGTGAAGACATGCCATTAAGCTT 3749
3724 AGCATGCTTAACGTGTAAGGCTGTCCACCACTGAAACATGCGGAGACATGTGTGCTATA 3783
3750 AGCATGTTGACCAAGAAAGCTTGTCTGCACTGTAATCCGCGGAACTGTGTGACGATA 3809
3784 GGTATGAGGCTTGTGATTCGGGCAACCGAATATCATCATCTGGGTGGGACGCTCATTT 3843
3810 GGTATGATGTTACGCTGACAGGCGGAGAAACATGATGTGTATTAAGCGGCGAGTTC 3869
3844 AGGTTACCGGTGTCTGTACAGCTTAAGAACCTGCCGAAATATCTGAGTCTCTGCTG 3903
3870 AAGTTTCCCGGATATGCAAAACGAATCTCATCTTGAAGAGAGAGTCTGTGTTGTA 3929
3904 TTTCTGGGAGAGCAACGCGCAACCAACATGACAGAGAGAGTCCAGAGCTTCAAGTGTG 3963
3930 TTTCAATGGGTGATGCCAAGGCGGTAAGGCAATCTTCAAGCTTTCATCAACCTTG 3989
3964 GACAACTATATCAAGGTGTAACAGATGAGAGGAGGAGAGTCCAGAGCTTCAAGAGTG 4023
3990 ACCAACTTTATTAAGGTTCCAGCTCCAGCAAGCCGATGTGACCTCTCATATCATGTG 4049
4024 ATCAGAGTGCATTTAGCAAGAGCGCTGACCAAGCTATCGTTAATGCTGCTAATAGCAAA 4083
4050 GTGGAAGGGAATATTTGCCAGGCAACCGAAGAGATGATTAATATGCTGTCTAAGCAAG 4109
4084 GGTCAACCAAGTTCGAGTGTGCGGTGCACTGTACGGAATGAGCCGCGCTCTTTGAT 4143
4110 GGAACAACCTGGCGAGGAGGTGTGCGGAGCGCTGTATTAAGAAATTCGCGAAGAGCTTGCAT 4169

QY	4144	AGACGCCAAATAGCTGTGGGACGGCTGACTGTGGAAACGAACCC---GCTATCTATA	4200
Db	4170	TTACAGCCGATTCGAGTAGAGAAAAGCCGCACTGTGCTAAAGGTGCACTAAACATATTCATT	4222
QY	4201	CATGCTGTAGGACCCCAATTTTTCTTAAGATGCCGGAACCGGAGGGCACTTAAAGCTCGCA	4266
Db	4230	CATGCCGTAGGACCAAACTTCAACAAAGTTTGGAGGTTGAAGGTGACAAACAGTTGGCA	4289
QY	4261	GCTGCTTCATAGAGATAGCCGTCCATTCGTCAACGCTGAGCCGATTAACAAAATATTCAGTA	4322
Db	4290	GAGGCTTATGAGTCCATGCTTAAGATTGTCAACGATAACATTAACAGTCAAGTACGATT	4345
QY	4321	CCGCTACTGTCAACCCGGCATCTATTCTGGTGGCAAGATTCGAGTATCAATTCATTCGCAT	4380
Db	4350	COACTGTGTCCACCCGGCATCTTTTCGGGAACAAAGATTCGACTAACCCATCAATTGAAC	4400
QY	4381	CACCTGTTCACTGCTTTTGACACTAGAGTCCGATGTCCATATATTTGCTGGATATA	4440
Db	4410	CATTGTCTGACAGCTTTTAGACCACTCATGTGAGATGTAGCCATATTCTGCGGGACAAG	4466
QY	4441	CAATGGGAGACCAAGGATATATCGAGGCCATTCAACCCAAAGAAAGGCTGGAATTTGGA-	4499
Db	4470	AAATGGGAAATGACTCTTCAAGGAGCAATGCTGTAGAGAGAAAGCATGGAGGAGATTATGC	4529
QY	4500	-----TGATGACAAAGCCAGTAGACATTGACTTGGTCAGGGTCCACCCAAAC	4545
Db	4530	ATTATCCGACGACTTTTCAGTACAGAAACCTGATGACAGCTGTGATGAGGGTGATCCGAAG	4589
QY	4546	AGCTTTTGGGAGAGACGACAGGTTATCTCCGTCAATGAGGGCAAGTTGATTTCAATCTG	4606
Db	4580	AGTTCTTTGGCTGGAAGAGAGGGCTTAGACACACAGACGATGGCAAAACTTTTCTCATATTTTG	4649
QY	4606	GAAGTACACGATTCATCAGACCGCCCAAGGACATTGGCCGAATTCATGACATGTGGCCC	4665
Db	4650	GAAAGGAAACCAAGTTTCAACGAGCGGCCCAAGGATATAGCAAAATTAATGGCCATGTGGCCC	4709
QY	4666	AACCAATTTGAGGCTTAATGAGCAAGTTTGTCTTGTACATCTGTGGGGAAGATATGTCCAGC	4722
Db	4710	GTTGCAACGGAGGCCAATGAGAGGATATGACATGATATCTTCGAGAAAGACATGTGACAGT	4763
QY	4726	ATCCGCTCCAATGGCCAGTAGAGAGATCAGAGGGCTGTGCTCCACCTCACACATTCCCA	4785
Db	4770	ATTAGTGTGAATGTCCCGCTCGAAGAGTGTGAAGGCTCTCACCACCTAGACAGCTGTGCTT	4822
QY	4786	TGCTGTGTATAATTACGCTATAGACGGCTAGAGCGCTATAACAGTTGGCTTGTCCGAAGAA	4845
Db	4830	TGCTTGTGTATCCATGCGCATGACTCCAAAGAAAGATACAGCGCTTAAAGCCTCAAGTCCA	4889
QY	4846	GAAACAGTTCCGGCGTATGTGCTCATATCTCTGTGGCCGAAGTACAGATTCACAGGCTGTGAG	4905
Db	4890	GAAACAAATTACTGTGTGTCTCATCTTTTCATTGTCGGAAGTATAGATATACGTGTGTGAG	4949
QY	4906	AAGCTACAGTCCAGCAAAACAGACTCTGTTTTTCAGGCTGTGATCAACACCGGCTGTACACCC	4965
Db	4950	AAAGATCCATGTCTCCACAGCTTATATGTTTTCACCGAAAGTGTCTGGCTATATTCATCTCA	5009
QY	4966	AGGAAGTA 4973	
Db	5010	AGGAAGTA 5017	

RESULT 6
AAD29132
ID AAD29132 standard; DNA; 7479 BP.
XX
AC AAD29132;
XX
DT 07-MAY-2002 (first entry)
XX
XX Venezuela equine encephalitis virus (VEE) non-structural gene.
XX
XX Venezuela equine encephalitis virus; VEE; vaccine; HIV infection.
KM

human immunodeficiency virus; alpha-virus replicon; immune response;
therapy; non-structural gene; ds.

Venezuelan equine encephalitis virus.

OS

XX Key Location/Qualifiers
FH 1..7479
FT CDS
TT /tag= a
CTT /product= "VEB non-structural protein"
TT /trans_except= (pos:5635..5643, aa:Gln-Arg)
TT /note= "insertion of 3 bases alters the reading frame;
TT CDS does not include stop codon"

Pt partial

FN MO200203917-A2.
EN
PD 17-JAN-2002.
PE 09-JUL-2001; 2001WO-US21701.
PR 07-JUN-2000; 2000US-216995P.
PX (ALPH-) ALPHAVAX INC.
PA (UYNC-) UNITV NORTH CAROLINA.
PI Olmsted R., Keith P., Dryga S., Caley I., Maughan M., Johnston R.;
PI Davis N., Swansstrom R;
PI
XX MPI; 2002-171664/22.
DR P-PADB; AAEI8302.
XX Composition useful for treating or preventing HIV infections, comprises
PT two or more isolated nucleic acids encoding env, gag or pol gene
PT product of HIV or immunogenic fragment of the gene products -
PS Example 2; Page 145-156; 201pp; English.

The invention relates to a composition comprising isolated nucleic acids
encoding env, gag or pol gene product of human immunodeficiency virus or
immunogenetic fragment of the gene products. The gag gene product is
modified to inhibit formation of virus-like particles containing gag
gene product and their release from cells, and the pol gene product is
modified to inhibit reverse transcriptase activity. The invention also
relates to a method for producing an alpha-virus replicon particle used
in vaccines. The composition is useful for inducing an immune response
to human immunodeficiency virus (HIV) or for treating or preventing HIV
infection in a subject. The alpha-virus replicon particle is useful in a
vaccine. The composition is useful for administering a protein or peptide
to a subject. A composition comprising heparin affinity-purified alpha-
virus replicon particle is useful as a clinical trial material and as a
commercial product. The present sequence is a non-structural gene from
Venezuelan equine encephalitis virus (VEE) used in the exemplification
of the invention.

SQ Sequence 7479 BP; 2141 A; 1774 C; 1900 G; 1664 T; 0 other;

XO Query Match 42.2%; Score 2111; DB 24; Length 7479;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3209; Conservative 0; Mismatches 1740; Indels 24; Gaps 3

OY 25 ATGAAGAATTCACGTGGACTTAGAGTGACACC GCCGTATGTCAGTCCGTTACAACGCC 84
DB 1 ATTGAGAAAGATTCACTGCATCGAGAMACAGC CATTCCTFAGAGCTTTGACAGCG 60
. . . AC GT TT CC ACA CT TG AG AT CGA GC CA GA GG AG CT ACT G CA AA TA GA ACC A TC GC 144
Db 61 AGCTTCCC GCA GTT G AG TGA AG AAC CA GA GG CA CT G AT GA A TA GA CC AT G CTA AT GC 120
145 AGAGCGTTTTTCGATGTGCAACAAGGCTCATTTAGAGCGGAAGTGCACCGGACCAAAGTT 204
Db 121 AGAGCGTTTTTCGATGTGCAACAAGGCTCATTTAGAGCGGAAGTGCACCGGACCAAAGTT 204
OY 205 ATCTTGACATTTGAGAGTGGCCCGCTCACAGATGCACATTTCCAATCACCGCTATATTGT 264

181 ATCTTGACATTTGAGGAGGCGCGCGGAGAAATGATTTCAAGCACAAGATTCATTGT 240
Qy 265 ATTCGCCCTATGATTAAGCGCTGAAGACCCGCGACAGATCAACCGTATGCGAAAGACTT 324
Db 241 ATCTGTCCGATGAGATGTCGGAAGATCCGACAGATGTGATTAAGTAACTTAAGCTG 300
Qy 325 AAGAAA-----AGTGAATTACCGGACAAAGAACATAGCTCTTAAGGCGGAGACCTGCTG 378
Db 301 AAGAAAACCTGTAAGAAATTAATGATTAAGAAATTGGAACAAGAAATGAAGAGCTCGCC 360
Qy 379 GAATGATGTCACACGACGACGACAGATCTCATCTGTGTATGACACACAGCGCCACG 438
Db 361 GCGGTCAAGAGCGACCTGACCTGGAACCTGAGACTATGTGCTCCACGACGAGAGTGC 420
Qy 439 TGTAGTACTTTGGAAGTGAAGTACAGATATCCAAATGTGTACGAGTCCATGACCGACA 498
Db 421 TGTGCTACGAAGGCGCAAGTGCCTGTTTACAGAGATGTATAGCGGTTTGAACGACGACA 480
Qy 499 TCAATCTACACGACGAGCGCTTAAGAGGTTAGGACAATTTAATGGAATGAGCTTTGACACG 558
Db 481 AGTCTCTATACCAAGCCAAATTAAGGAGTTAGAGTGCCTTACTGATAGGCTTTGACACG 540
Qy 559 ACCCTTTTATGTAACAATAAGATGAGTTCCTACCTTATCAACACGAGTGGGCT. 618
Db 541 ACCCTTTTATGTTAAGAACTTGGCTGAGCATATCATCATATCTCAACGACGAGGCGC 600
Qy 619 GACGAGAGATTTGGAAGACGTAACATTTGGCTCGGTAACTGAGATCTTCAGAGAGAC 678
Db 601 GACCAAAACCGTGTACCGCTCGTAAACATAGCCCTATGCACTGACGTATGAGAGCG 660
Qy 679 AGGCTTGAAGAACTCTCAATCCTTAGGAGAAAGAGGCTCCAACTCATATTAAGTACA 738
Db 661 TCAGCTAGAGGAGATGCTCATTTCTTAAGAAAGATTTGAACCATCCAAAGTTCTTA 720
Qy 739 TTCTCGTGTGTTCAACATCTACACAGAGATAGATCACTGTAACTGAGTGCATCTT 798
Db 721 TTCTGTGTGCTGACCATCTACACAGAAAGAGGACTTACTGAGAGAGCTGGACCTG 780
Qy 799 CCAAAAGTGTCCACTTGAAAGAAAGTCTAATCTCAGAGTAAATGTGGGACCAATTGTC 858
Db 781 CCGTCTGATTTCACTTACGTGGGACAGCAAAATTAACATGTCCGTGTGACATTAAGTT 840
Qy 859 AGCTGTGAAGGAGTGCATCAAAAGATTAAGATGAGCCAGGACTTAAGGTTAAAGTT 918
Db 841 AGTTGCGAGGAGTACGTCTTAAGAAAGATTAAGTATCAAGGCTGTATGAGGAGCTT 900
Qy 919 GAGAACCTTGGCGTCCACATGCAATCGCAGAGGTTCTTGAAGTTGCAAAATCAAGATACG 978
Db 901 TCAGGCTATGCTGCTACGATGCAACGCGAGGGAATCTTGTGCTGCAAAAGTACAGACA 960
Qy 979 CTGCGCGCGAGAGGCTTTCTTTGCTGTGTGTAAGTATGTAACAGACCACTTTGCGAT 1038
Db 961 TTAAACGGGAGAGGAGTCTTTTCCCTGTGCAAGTATGCAAGCTTGTGTAC 1020
Qy 1039 CAGATGACAGGAGTTCTGGAAGTGAAGTGTGATGTAAGTGAAGTGAAGTGAAGTGAAGTGA 1098
Db 1021 CAAATGATGCGATATGCGCAACAGATGTCAAGTGTGAGGAGCAGCGCAAAATCTGCTGTT 1080
Qy 1099 GGGCTCAACCAAGAGATGTGCTCAATGTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1158
Db 1081 GGGCTCAACCAAGATATGCTGCAACGCTGCAACCCAGAAACCAACATACATGA 1140
Qy 1159 AACTATCTATTACAGATGCTGCGCCAGCGCTTTTCCAGGTGGCGCGTGAACATGCTGCC 1218
Db 1141 AATTACCTTTTCCCGTGTGCTGCGCCAGCATTTGTGTAGGTGGGCAAGAAATATTAAGAA 1200
Qy 1219 GACTTGAAGCAGAGAAAGAACTAGGGGTGGGGAGCGCACTCTTAATAGGCTGCTGCC 1278
Db 1201 GATCAAGAGATGAAGGCGCACTAGGACTAGAGATAGACAGTTATGTCAGGGGTGTTGT 1260
Qy 1279 TGGGCTTTCAAGACCCAGAAATCAATTCATCTTACAGAAAGCCTGTGATCCAAATTT 1338
|||||

1261 TGGGCTTTAGAAAGCACAAGATTAATCTATTATTAAGGCGCGGATACCAAAACATC 1320
Qy 1339 AAGAAATACCTGCGCTTTGACTCATTTGTATTCACAGCCCTTACAGCCACAGGCTC 1398
Db 1321 ATCAAGTGAACAGCGATTTCTCACTATGCTGTGCTCCAGAGATAGGACATGAATG 1380
Qy 1399 GATATGGGCTTCCGCGCTTGAAGCTCAAGCTGCTGTTAAACCACTGTCAAAACCCGACCG 1458
Db 1381 GAGATCGGCTGAGAACAAAGATCAGAAATGTTAAGAGAGCACAAGAGCCGCTACCT 1440
Qy 1459 GCTATTAACAATGCGGATGTGAGCATCTGCGGTCTTTACAGCAAGAGCTGAAGAAATG 1518
Db 1441 CTCATTACCCCGGAGAGTGAAGAAAGTAAAGTCCAGCGATGAGGCTTAAGAGAGT 1500
Qy 1519 GCTGCAAGGAGAGATCAGAAAGCCCTGCAACCTTGTCTCCCTGAATTAAGAAAGAG 1578
Db 1501 GGTAAAGCCAGAGGTTGCGGCGAGCTTACCACTTTTGGCAGCTGATGTGAAGAGCC 1560
Qy 1579 ACCGTAGAGCAGAGATGACCTCATTTATGCAAGAGCAGAGCAGGTAGCGTGAACA 1638
Db 1561 ACTGTGAAGCCGATGTCGATTTGATGTTACAAAGGCTGGGCGCTCAGTGAAGACA 1620
Qy 1639 CACGAGAGACATCAGGGTGAAGTTACCAGGCGGAGAGAAAGATTTGGCTTTACGCT 1698
Db 1621 CCTGCTGGCTGATTAAGATTTACAGCTACGCTGGCAGAGACAAGATCGGCTTTACGCT 1680
Qy 1699 ATACTTTCAACCCGAGGCTATTTGAATAGTGAAGAAACTGGCGGTATCCACCCATTGGCG 1758
Db 1681 GTGCTTTCTCCGAGGCTGTATCTCAAGTGAAGAAATTTATCTTGCAATCCCTCTGCT 1740
Qy 1759 GAACAACTACTGTATGACTCAAAAGTGAAGGCGAGGATCAAAAGTGCAGCCATAC 1818
Db 1741 GAACAACTACTGTATGACTCAAAAGTGAAGGCGAGGATCAAAAGTGCAGCCATAC 1800
Qy 1819 CACGTAAGCTATTTGAACAGAAAGGACGCGGCTCCCTGTTCAAGCTTCAGGCAATTG 1878
Db 1801 CATGTAAAGTATGAGTGCAGAGAGGACATGCAATACCCGCTCCAGGACTTTCAAGCTCTG 1860
Qy 1879 AGTGAAGGCTACGATCGTTTCAAGAGAGGAGTTCTGAACAAATCTGACACAC 1938
Db 1861 AGTGAAGGCTACGATCGTTTCAAGAGAGGAGTTCTGAACAAATCTGACACAC 1920
Qy 1939 ATCGCAATCAACGAGAGGCGCTTAACACTGACAGAAAGTACTTAAGCTGTAAAGACT 1998
Db 1921 ATTGCCACATGAGAGGAGCGCTGAACATGATGAAGATTAACAAACCTGTCAACCC 1980
Qy 1999 CAGGACACAGCTCAAGATACGTTCTGATTTGAATGACGACGAAAGTGTGAAGCGAGAA 2058
Db 1981 AGCAGACACAGCGCGAATACCTGTACGACATCGACAGAAACAGTCCGTCAAGAAAGAA 2040
Qy 2059 GAGCGAGTCCCTGTGCTTAACCGGTATCTGTGATCAACCACTTCAAGCTTGGCG 2118
Db 2041 CTATGCTAGGCTGAGGCTTCAAGGAGCTGAGTGTGATCTCTTCAAGATTTGCC 2100
Qy 2119 TACGAGAGTCTCAAGACAGACAGCAGACCTTCAACAAAGTCCCAACCATGAGTAT 2178
Db 2101 TACGAGAGTCTCAAGACAGACAGCAGCCTCTTCAACAAAGTCAACCATGAGGCTAT 2160
Qy 2179 GAGTGCAGGTTCAAGTAAATCTGAATCATCAAAAGCGCTGTGACTTAAGAAAGATCTG 2238
Db 2161 GCGCTGCAGAGATCAGCAAGTCTGCGATCATTTAAAGCGCAGTCAACAAAAGATCTA 2220
Qy 2239 GTTGTAGTGCAGAAAGAAACTGGGCAAGAAATCACTAGGATGTAAGAGAGATGAGA 2298
Db 2221 GTGTGAGCGCGCAAGAAAGAACTGTGCAAGAAATTTAAGGAGCTCAAGAAATATGAA 2280
Qy 2299 GGTATGATGTGTGCTAGAGACTGTGATTCAGTGTCTTCAATAGGGTTAAGCAACCC 2358
Db 2281 GGGCTGACGTCATGTCAGAACTGTGACTGAGTGTCTTGAATGATGAAACACACCC 2340
Qy 2359 GTTAACACTCTGTATCATTTGATGAAGCAATTTGCTGCAATGAGGAGCGTGTGCACTG 2418
Db 2341 GTAGAGACCCGTATATTTAGAGAAAGCTTTGTGCTTGTATGACGAGTACTTCAGAGCGCTC 2400
|||||

QY 2419 ATTGCATCGTCAACCTTAAGAAAGTGGTATTTGTCGGGAGCCCAAAACAAATGCGGCTTC 2478
 DB 2401 ATAGCATTTATTAAGACTTAAGAGAGAGTGGCTCTGCGGGATCCCAACAGTGGCTTTT 2460
 QY 2479 TTTTAACATGATGTGCTGTAAGTACATTTTAAACATGACATATGACATGAGTACAT 2538
 DB 2461 TTTTAACATGATGTGCTGTAAGTACATTTTAAACATGAGATTTGACACAAATCTTCCAC 2520
 QY 2539 AAAAGCATCTCTAGAGGAGTGCACAGACTGTAACCGGCATGTCTCAACGCTCTTCCAC 2598
 DB 2521 AAAAGCATCTCTGCGCTTGCATTAATCTGACTTGGTGTCTCAACCTTGTTTAC 2580
 QY 2599 GACACGCAATGAAGACGTTTAAACCATGTGCTGTAATAATCATATGATATACACAGG 2658
 DB 2581 GACAAAAAATGAGAACGACGATCCGAAGAGATTAAGATTGTATGACACTACCGG 2640
 QY 2659 ACCACAAACCGGACAAAGATGATCTGATTTTCACTGTTTCAAGATGGGTGAACAG 2718
 DB 2641 AGTACCAACCTTAAGACAGACGATCTCATCTTGTTCAGAGGGTGGTGAACGAG 2700
 QY 2719 CTACGATTTGACTCAAAATATCAGAAATCATGAGGGGCTGATCCGACAGGATTAACG 2778
 DB 2701 TTGCAATTAATGATTAAGAGGACGAAATTAATGACGAGCTGCTCTCAAGGGCTGAC 2760
 QY 2779 CGAAAGGCGTTTATGCTGTCAAGTCAACAGGAAATCACTCTACTCGCAGCT 2838
 DB 2761 CGTAAAGGTGTATGCGCTTGTGTAAGAGTGAATGAATAATCTCTGTATGACACCCAC 2820
 QY 2839 TCTGACAGTGAACGTGTACTTAACGACGACAGAAAAACGATGTCTGAAAGAGCTA 2898
 DB 2821 TCAGAACATGTGAACGTCTCTACTGACCGGACGAGAGCCGATGTGTGAATAACATCA 2880
 QY 2899 GCTGTGATCCCTGATTAAGACACTTAACGTAATATCCGGGATTTTCAAGCTTCA 2958
 DB 2881 GCCGCGACCAATGATTAAGAACTGACGCAAGTACCTCGGAAATTTTCACTGCCAG 2940
 QY 2959 TTGACGACTGGACGCGGAAACGACGCAATTAATGACGCGCTTGTATTAAGCGAG 3018
 DB 2941 ATAGAGAGTGGCAAGACAGATGATGATCATATGAGGACATCTTGAAGAGACGAG 3000
 QY 3019 ACAGCTGATGTGTTCCAGAAATTAAGTGAACGTCTGCGGCGAAAGCTTTTGAAGCAGTC 3078
 DB 3001 CCTACCGAGCTTCTCAAAATTAAGGCAACGTTGTGGGCAAGGCTTTTGTGCGGCTG 3060
 QY 3079 TTGGCAGGCGCAATTTGTGCTGACGAGACAGAGTGGAGACCTTGAACCATTCAG 3138
 DB 3061 CTGAAGACCGCTGGCATGACATGACCACTGAACATGGAACACTGTGATTTATTTGAA 3120
 QY 3139 CATGACAGAGCTACTCACTGAATGACACTGAATCTTTTGCACAGGTTCTTTGGA 3198
 DB 3121 ACGGACAAAGCTCACTACGACAGATGATTAATGAACCAACTATGCTGAAGGTTCTTTGGA 3180
 QY 3199 GTAACCTGGAACAGTGGTTATTTTCCGCTCTACCGTGCACATTACTTAACAGGATCAG 3258
 DB 3181 CTGATCTGGAACCGGCTCTATTTTCTGACCCCACTGTTCGGTTATCCATTAGGAATAT 3240
 QY 3259 CACTGGGATTAATCGCCAGGGAAGACATGATGGGCTTTATAGAGATGCAAAAGAG 3318
 DB 3241 CACTGGGATTAATCGCCAGTGGCTTAACATGATGAGGCTGAATTAAGAGATGCTCCGTCAG 3300
 QY 3319 TTGTACAGGCGATATCCGTGATCACAAAAGCGTTGACACAGGAGGATGATGATA 3378
 DB 3301 CTCTCTGCGAGTACCCCAAACTGCTCGGCGAGTTGCACTGGAAGGATCTATGATACATG 3360
 QY 3379 AGGAATTAATACATCAAGAGACTCTTCCAAATTAATGTTGTTCCATTAATGCGCG 3438
 DB 3361 AACACTGTATACACTCGCAATTAATGATCGGCGAATAAATCTAGTACCTGTAACGAGA 3420
 QY 3439 TTGCGCCACTGTTGATGTTGACCAAAAGGACGGGTACAATGATCAACAGCGATTC 3498
 DB 3421 CTGCTCATGCTTTAGTCTCTCCACCATATGAAACCCACAGAGTACTTTTCTTCATTC 3480

QY 3499 CTATCTAAGATGAAGGCGAAATCTGTGTGATGATCGGCGATCTATCAGATTCAGAG 3558
 DB 3481 GTCAAGAAATGAAGGGGAGAACTGTCTGTGTGTCGGGGAAGATGTCCGTCCAGAC 3540
 QY 3559 AAGAAATGAGTCCATGAGGTTCATTTGCCACTTAATACATCAGGTGTGATCTGATTTG 3618
 DB 3541 AAAATGTTGACTGTTGTGACACCGGCTAGGCTACTTCAAGGCTCGGCTGATTTA 3600
 QY 3619 GGAATACCTAGCCATGTCGTAATATGACATTAATCTTTGCAATGTTAGGACCCCGGTC 3678
 DB 3601 GGCATCCAGGTGATGTGCCAAATATGACATTAATTTGTTAATGTAGAGACCCCATAT 3660
 QY 3679 AGGAACATCACTACCAACAGTGCAGAGATCAAGCTATCAACACAGCATCTACGTGT 3738
 DB 3661 AATATCATCACTATACAGAGATGTAAGACCATGCAATTAAGCTTAGCATGTGACCAAG 3720
 QY 3739 AAGGTGTCCACCACTGTAACACTGGCGGAACATGTGTGCTATAGGATAGGCTTGTCT 3798
 DB 3721 AAAGCTGTCTGCACTGTGAATCCGCGGGAACCTGTGTACAGATAGGTTATGTACGCT 3780
 QY 3799 GATGCGCAACCGAATATCATCATCTGCGGTGGACGCTCATTTAGGTTTACCCTGTC 3858
 DB 3781 GACAGGCGCAGGAAAGCATATGTGTCTATAGCGCGGACGTTCAAGTTTCCGGGTA 3840
 QY 3859 TGTGAGCTTAAGAACACTGCGGAAATTAAGAGTTCTTCTGTGTTCTTCCGCAAGAC 3918
 DB 3841 TGCAACCGAAATCTCACTTGAAGAGACGAAATTCGTTTGTATTCATTGTTGGATGAT 3900
 QY 3919 AACGCAACCAACATGACCAAGACAGACTCGGTGATGTGCTTGAACATCTATCA 3978
 DB 3901 CGCAAGGCGCGTACGCAATCTTACCAAGCTTTCATCAACCTTGACCAATATTAACA 3960
 QY 3979 GGGTCAACCAAGTACGAGGAGGAGAGTCCAGGCTCAAGATGATCAGAGTACAGATT 4038
 DB 3961 GGTTCAGACTTCACAGAGCGGATGTGACCTTCATATCATGTGTGCGAGGGATAT 4020
 QY 4039 AGCAAGAGCGCTGACCAAGCTATGTTAATGCTCTAATAGCAAAAGTCAACAGATTCC 4098
 DB 4021 GCCACGCGCAACGAAAGATGATTAATATCTCTGTAACACAAAGACCAACCTGCGGA 4080
 QY 4099 GGAAGTGTGCGTGAATGTCGGAATATGCGGCTGCTTTGATAGACGCCAATAGCT 4158
 DB 4081 GGGGTGTGCGGAGCGCTGATTAAGAAAGTCCCGGAAAGCTTCGATTTACACCGCATCA 4140
 QY 4159 GTGCGACGCTAATCTTGTGAACACGAAAC---GCTCATCATATCATGCTGTGAGACC 4215
 DB 4141 GTAGAAAAAGCGCACTGTGTAAAGTGCACATTAACATATTCATGCGGTAGGACCA 4200
 QY 4216 AATTTTCTAAGATGCGGAACCGGAGGCGACCTTAAGCTCGCAGCTGCTACATGAGC 4275
 DB 4201 AACTTCAACAAAGTTTGGAGGTTGAAGGTGAACAAACAGTTGGCAGAGCTTATGAGTCC 4260
 QY 4276 ATAGCGTCACTGTCACGCTGACGGATTAACAAAATATCAGTACCGCTACTGTCAAC 4335
 DB 4261 ATGCTAAGATTTGCAAGATTAACAATTAACAAGTACAGATTCACCTGTGTCACC 4320
 QY 4336 GGCATCTATTTGTGTGCAAAAGATGCAATGATGCAATCATCTGCTTCACTGCT 4395
 DB 4321 GGCATCTTTTCCGGAAACAAAGATGCACTAACCAATCATTTGAACCATTTCTACAGCT 4380
 QY 4396 TTGACACTAGATGCGATGTCAACATATATGCTTGTGATTAACAATGGGAGACGAG 4455
 DB 4381 TTAACACACTGATGACAGATGTAGCCATATATCTGAGGACCAAGAAATGGAAATGACT 4440
 QY 4456 ATATGAGGCAATTCACCGCAAGAAAGCGTCAAAATTCGGA-----T 4500
 DB 4441 CTCAAGGAGAGAGTGTAGGAGAGAGCAAGTGAAGAGATATGATTCGACGACTCT 4500
 QY 4501 GATGACAGCGATGATGATCTTGTGCAAGGCTTCAACCAACAGCTCTTTGGCAGGC 4560
 DB 4501 TCAATGACAGAACTGTATGACAGAGCTGTGAGGTGATCGAAGAGTTCTTTGGCTGGA 4560
 QY 4561 AGACAGGTTACTCGTCAATGAGGCGAAGTTATTCATATCTGGAAGGATACAGATTC 4620

Db	461	AGGAAGGCTACAGCAGACGCGTGGCAAACTTCTCATTTTGGAAAGGACCAAGTTT	4620
Qy	4621	CATCAGACCGGCCAAGACATTGGCCGAATCCATGCAATGTGGCCCAACAAATCTGAGCT	4680
Db	4621	CACACGCGGCCAAGAGATATAGCAGAAATTAATGCCATGTGGCCCGTTGCACCGAGGCC	4680
Qy	4681	AATGAGCAGATTTTCTGTATACATCCTGGGGAGAGTATGTCACAGATCCGCTCCAAATGC	4740
Db	4681	AATGAGCAGATATCATATATATCTCGAAGAAAGCAGACAGATTAAGTCCGAATGC	4740
Qy	4741	CCAGTAAAGAGTCAGAGGCGTCTGCTCCACCTCACAACATTCATGCTGTGTAATTAC	4800
Db	4741	CCCGTCGAAGAGTCGAAGCGCTCCACACCACTTAGACAGCTGCTTGTGTTGATCAT	4800
Qy	4801	GCTATGACGCGTGAAGCGCGTATTAACAGTTGGCGCTCTGCGAAGAAAGAACAGTTGCCGTA	4860
Db	4801	GCCATGTACTCCAGAAAGAGTACAGCGCCTTAAGACCTCACGTCGAAGAACAAATTACTGTG	4860
Qy	4861	TGCTCATCATTCCTGTTGCCGAAGTACAGAGATCACAGGCGTCGCAAGAGTACAGTGCAGC	4920
Db	4861	TGCTCATCTTTTCATATGCCAGAGATATGAATCACTGGTGGCAGAAATCAATATGCTCC	4920
Qy	4921	AAACGAGTCCTGTTTTCAGGCGGTGATCCACGCGGTGATCACCCGAGAACTTA	4973
Db	4921	CAGCCTAATATGTTCTCACCGAAAGTGTCTCGATATATTCATCCAGGAAGTA	4973

CC	XX	DR	XX	PA	(REED-) REED ARMY INST RES WALTER.
CC	XX	PI	XX	PB	Crisse BJ, Oberste MS, Parker MD, Schumura SM, Smith JF;
CC	XX	PT	XX	PF	WPI; 1999-045316/04..
CC	XX	PS	XX	PR	New DNA encoding infectious Western or Venezuelan equine
CC	XX	RD	XX	PP	encephalitis virus genome - useful for the production of live or
CC	XX	RE	XX	PD	attenuated vaccines for human or veterinary medicine
CC	XX	RH	XX	PE	
CC	XX	RI	XX	PG	
CC	XX	RJ	XX	PH	
CC	XX	RK	XX	PI	
CC	XX	RL	XX	PJ	
CC	XX	RM	XX	PK	
CC	XX	RN	XX	PL	
CC	XX	RO	XX	PM	
CC	XX	RP	XX	PN	
CC	XX	RS	XX	PO	
CC	XX	RT	XX	PP	
CC	XX	RU	XX	PR	
CC	XX	RV	XX	PS	
CC	XX	RW	XX	PT	
CC	XX	RY	XX	PU	
CC	XX	RZ	XX	PV	
CC	XX	SA	XX	PW	
CC	XX	SB	XX	PX	
CC	XX	SC	XX	PY	
CC	XX	SD	XX	PZ	
CC	XX	SE	XX	QA	
CC	XX	SF	XX	QB	
CC	XX	SG	XX	QC	
CC	XX	SH	XX	QD	
CC	XX	SI	XX	QE	
CC	XX	SJ	XX	QF	
CC	XX	SK	XX	QG	
CC	XX	SL	XX	QH	
CC	XX	SM	XX	QI	
CC	XX	SN	XX	QJ	
CC	XX	SO	XX	QK	
CC	XX	SP	XX	QL	
CC	XX	SQ	XX	QM	
CC	XX	SR	XX	QN	
CC	XX	SS	XX	QO	
CC	XX	ST	XX	QP	
CC	XX	SU	XX	QQ	
CC	XX	SV	XX	QR	
CC	XX	SW	XX	QS	
CC	XX	SX	XX	QT	
CC	XX	SY	XX	QU	
CC	XX	SZ	XX	QV	
CC	XX	TA	XX	QW	
CC	XX	TB	XX	QX	
CC	XX	TC	XX	QY	
CC	XX	TD	XX	QZ	
CC	XX	TE	XX	RA	
CC	XX	TF	XX	RB	
CC	XX	TG	XX	RC	
CC	XX	TH	XX	RD	
CC	XX	TI	XX	RE	
CC	XX	TJ	XX	RF	
CC	XX	TK	XX	RG	
CC	XX	TL	XX	RH	
CC	XX	TM	XX	RI	
CC	XX	TN	XX	RJ	
CC	XX	TO	XX	RK	
CC	XX	TP	XX	RL	
CC	XX	TQ	XX	RM	
CC	XX	TR	XX	RN	
CC	XX	TS	XX	RO	
CC	XX	TT	XX	RP	
CC	XX	TU	XX	RR	
CC	XX	TV	XX	RS	
CC	XX	TW	XX	RT	
CC	XX	TX	XX	RU	
CC	XX	TY	XX	RV	
CC	XX	TZ	XX	RW	
CC	XX	UA	XX	RX	
CC	XX	UB	XX	RY	
CC	XX	UC	XX	RZ	
CC	XX	UD	XX	SA	
CC	XX	UE	XX	SB	
CC	XX	UF	XX	SC	
CC	XX	UG	XX	SD	
CC	XX	UH	XX	SE	
CC	XX	UI	XX	SF	
CC	XX	UJ	XX	SG	
CC	XX	UK	XX	SH	
CC	XX	UL	XX	SI	
CC	XX	UM	XX	SJ	
CC	XX	UN	XX	SK	
CC	XX	UU	XX	SL	
CC	XX	UV	XX	SM	
CC	XX	UW	XX	SN	
CC	XX	UX	XX	SO	
CC	XX	UY	XX	SP	
CC	XX	UZ	XX	SQ	
CC	XX	VA	XX	SR	
CC	XX	VB	XX	ST	
CC	XX	VC	XX	SU	
CC	XX	VD	XX	SV	
CC	XX	VE	XX	SW	
CC	XX	VF	XX	SX	
CC	XX	VG	XX	SY	
CC	XX	VH	XX	SZ	
CC	XX				

CC encephalitis (VEE) viral genome; infectious or attenuated VEE RNA
CC transcripts and viral particles; an attenuated chimeric virus (av)
CC containing non-structural sequences from a first alpha-virus (av)
CC and structural sequences from a second av, resulting in attenuation
CC of the second av; a method for expressing a protein (especially an
CC antigen to protect against a pathogen) by cloning its gene into an
CC attenuated VEE or VEE IE replicon so that transcription of the
CC replicon produces RNA able to infect the cells in which protein is
CC to be produced; methods for diagnosing VEE and VEE infections; and
CC polypeptides encoded by VEE variant IIA (see AA174109). Attenuated
CC VEE and VEE are used in live or inactivated vaccines, for use in
CC human or veterinary medicine. Chimeric viruses are also useful as
CC vaccines, directed against the second av which is particularly
CC Eastern equine encephalitis (EEE) virus or some variant of VEE.
CC VEE and VEE nucleic acids are used as primers and probes to
CC diagnose virus infections and to define natural variants, also for
CC production of protein antigens which can be used as diagnostic
CC reagents, to generate antibodies, and in vaccines. The attenuated
CC viruses are highly immunogenic and provide long-lasting protection.
XX
XX
Sequence 11464 BP; 3259 A; 2734 C; 2900 G; 2571 T; 0 other;

Query Match 41.5%; Score 2074.8; DB 20; Length 11464;
Best Local Similarity 64.1%; Pred. No. 0;
Matches 3202; Conservative 0; Mismatches 1772; Indels 24; Gaps 4;

Query Match	Similarity	Score	DB	Length
Best Local	64.1%	Pred. No. 0	Mismatches 1772	Indels 24
Matches 3202	Conservative	0		Gaps 4
QY	1	ACCCTACAAACCTAATCGATCCAAATGTGAAAGAAATTCACGTTAGATTACATGCTGACAC	60	
DB	20	AGCCCAACCAATATACTAACCCAAATGAGAAAGTTCAAGTTGACATGACGAAAGTAAGT	79	
QY	61	CCGTAATGCAAGTCTTACAGCGACGTTTCCACAATTTAGATGGAAGCAAGCGAGTTC	120	
DB	80	CCCTTCCCTCAAGACATTTACAAAGGAGCTTCCCGAGTTGAGAGTGAAGCCAAAGCAGGTC	139	
QY	121	ACTGACATGACCAATGCCAATGCCAGAGCGTTTGGCATGTGGCAACAAAGCTCATTGAG	180	
DB	140	ACAGTAATGACCATGCTACGACCGACAGAGGTTTTCGATTTGGCATTCGAATTTGATCGAG	199	
QY	181	AGCGAAGTCGACCGGAGCCAAAGTTATCTTGGACATTTGGAAGTGCGCCCGCTCAGACATGCA	240	
DB	200	ACGGAGGTGGAACCATTCGGATTCGATCTCTAAGCATTTGGAAGTGGCGCTGCCGACAAATG	259	
QY	241	CATTCCAAATCACCGCTATCATTTGTAATCTGCCCCATGATTAAGCGCTGGAAGACCCGACAGA	300	
DB	260	TATTCCAAGCATTAAGTACCATTTGCATCTGTGCGATGGAATGCGAGAAATCCGGACAGA	319	
QY	301	CTACAAAGCGTATGCGAAGAAAGCTTAAAGAAAGT-----GACATTACCGACAAGAACATA	354	
DB	320	CTGTTTAATGTAATGACGCCAAGCTGTAAGAAAGACTGTAAGAAATTAACAGATTAAGAACTG	379	
QY	355	GCCCTTAAAGCGCGAGACCTGCTGGAAGTCATGTCAACCCAGACGACAGACTCCATCT	414	
DB	380	GACAAAGAAATGAAGAGACTTGGCGAAGTCAAGACGACCTCGATCTGAAACTGAAACG	439	
QY	415	CTGTGTATGCAACAGACGCCACGTGTAGGTACTTTGGAAGTGTACAGTATACCAAGT	474	
DB	440	ATTGGCTTTCAGACGATGAACCTGTGATTTGAGGGTCACAGTGCACGTGATACGAGAT	499	
QY	475	GTTGACGAGTCCAGTACCGGACATCAATCTACACACAGGCGCTTAAAGAGTTAGAGACA	534	
DB	500	GTTGACGAGTTCGAGACGACCGACGACCTTTTACATCAAGGCCAACAAAGGGGTCAAGTTC	559	
QY	535	ATTTACTGATAGGCTTTGACACGACCCCTTTATGTACAAAAAATGACAGTTTCTTAC	594	
DB	560	GCTATTGGAATAGGATTGCAACCTACCCCTTTATGTTTAAAGACCTGCTGGAGCATAT	619	
QY	595	CTTACTTACAAACGAACTGGGCTGACAGAGAGTATTGGAAGCAGTAACTTTGGCCTTC	654	
DB	620	CCCTCTCTATTGACCAACTGGGCGCAGACGACGCTGTAAACGGCTGTATATATAGGCTTG	679	
QY	655	GGTAATCGAATCTTCAAGAGACGAGGCTTGGAAAACTCTCAATCCTTTGGAAGAAAGAG	714	
DB	680	TGCAGCTCCGATGTATGAGAGGCTCACTGTAAGGAGATGTCATCTCTCAAAAGAAATTT	739	

QY 715 CTCAACTACTAATAGATCATATTCGTGGTTCACAATCTACAGAGAATGA 774
DB 740 TTAATAACCAATCAATACGTCTGTTCTCTGTAGATCTACATCTACACAGAGAAGCA 799
QY 775 TCACGTTCAGTAGCTGGCATTTCCAAAGTGTTCACCTGAAAGAAAGTCTAATTC 834
DB 800 GACTTACTAAGAGATGGGACCTACCGTCCGTTTTCACCTACGTGTAAAGCAATTAC 859
QY 835 ACAGGTAGATGTGGACCATTTGTCAGCTGTGAAGGTACGTATCAAAAAGTAAAGATC 894
DB 860 ACTTCTGGTGTGAACCTATAGTTAGTTGCGACGGGTACGTGTCAAAAAGTACTATT 919
QY 895 AGCCAGAGACTATACGTAAAGTTGAGAACTTGCGCTCCACAATGCATCGAGGGTTTC 954
DB 920 AGTCAGGTCTCTACGAGGAAACCGTCCGGCTACGTCGCACAGATCGATCGAGGATTC 979
QY 955 TTGAGTTCGAAGTCACAGATACGCTGCGCGGAGAGGGTTCTTTTGTGTGTACG 1014
DB 980 TTGTCTCCAAAGTACGCGACACGCTTGACGGGAGAGGGTCTTTTCCCGTATGACG 1039
QY 1015 TATGTACAGCCACCTTTGCGATCAGATGACAGGATTTTGGCACTGACGTAGTGTG 1074
DB 1040 TACGTCCAGCCACATGTGTGTCGCAAAATGACAGGCAATTCGCAACAGATGACGTCA 1099
QY 1075 GATGACGACAAAACCTATTTGGTGGCTCAACCAAGATTTGTGTCATGTAGAGCG 1134
DB 1100 GACGACGCGCAAAATTTGCTGTGTGGCTCAACCAAGCAATGTGTAAATGTCGCGACT 1159
QY 1135 CAAGAATACTATACACATGACAGATCTATCTATTTACAGTGTGCGCCAGCGCTTTTC 1194
DB 1160 CAAGAATACTATACACATGACAGATCTATCTATTTACAGTGTGCGCCAGCATTTGTC 1219
QY 1195 AGGTGGGCGCGGAAATCTGTGCGCACTTTGAGACGACAGAAAGAACTAGGGTGGCGAG 1254
DB 1220 AGATGTGGCTAAAGATGACAGAAATGACAAAGATGAAAGCAATGGGGCTTGGGAC 1279
QY 1255 CGCACTCTTACTATAGGGCTGTGCTGGCTTTCAAGACCCAGAAATCATCATCTAC 1314
DB 1280 CGCCAGTGTGTATGGGGTGTGTGGCGCTTCAAGAAACACAAATCATCATCTGATAC 1339
QY 1315 AAGAACTGTGTACCAAAACAATTAAGAAAGTACCTGCGCTTTGACTCATTTGTGAT 1374
DB 1340 AAACGACGACACCCAAACGATCATCAAGGTAAACAGGATTCACCTTTGCTGTGCG 1399
QY 1375 CCAGCCTTACGACGAGGCTGATATGGGCTTCCGCGTACAGTCTGCTGCTT 1434
DB 1400 CCCAATTTGAGACCACTTTGAGATTTGGCTGAGGACCAAGATCAGAAACTACTG 1459
QY 1435 GAACCACTGTCAAAACCGGACCGGCTATTACATGCGCATGTGAGATCTGGGTGC 1494
DB 1460 GAGGAACCTGTGACAGACACCATGATTAACCGCGGACGATACAGAGAACCAAGAC 1519
QY 1495 TTAACGACAAAGCTGAGAAAGTGTGCTGACGCGAAGAGATCAGAGAACCCCTGCCACC 1554
DB 1520 GCGGGGATGTAGGCTTAAGGAAGTCAAGAGGCGAAGAGCTCAGGGCAGCATATCACACG 1579
QY 1555 TTGCTCCTGAATTAAGAAAAGACCGTAGAGGAGAAAGTACCTCATTAATGCAAG 1614
DB 1580 CTGTCTGCGATGTAGAAACCTGACCTGAGGCGGAGGTTGACTTAATGCTCAGAG 1639
QY 1615 GCAGAGAGAGGTAGCGTGAAGACACAGAGGACCATCAGGGTACAAAGTTACCCAGGC 1674
DB 1640 GCGGAGAGAGATCTGTGAAACCCCGAGGGGCTCATTAAGTCAACAGATTATCAGAGA 1699
QY 1675 GAAGAAGATTTGGTCTTAACCTTATCTTACCCGAGCGGTATTGAATAGTAAAAA 1734
DB 1700 GAAGCAAAATTTGCTTTATGCGGTGCTCTCCACAAGCAGTACTCGAAGTAAAAA 1759
QY 1735 CTGCGGTATTCACCACTTGGCGGAACAGTACTGTGTAATGATCTACAAAGGTAGGCA 1794
DB 1760 CTGAGGTCACTCAACCCCTTGCAAGAGCAAGTAATGTAATCAACACTCTGGAAGGAG 1819

QY 1795 GGGAGATCAAAAGTCAGACCATACCAAGGTAAAGTCATTGTACAGAAAGGACGCGTCC 1854
DB 1820 GGTAGATTCGAGAGTTGACCTTTACACGGAAGAAAGTGTAGTACCGGAAGGCAACTATA 1879
QY 1855 CTTGTTCAAGACTTTCAGGCACTTGAAGAGCGCTACGATCGTTTTCAACGAGAGGAG 1914
DB 1880 CTTGTCAAAGACTTTCAGGCACTTGAAGAGCGCCACAAATTTGTATCAACGAAAGTGTAG 1939
QY 1915 TTGCTTAACAGATACCTGACCAACATTCGCATCAACGAGAGAGCCCTAAACCTACGA 1974
DB 1940 TTGCTTAAGATTAATTAACCAACATTCGACGACGAGTGTAGAGCTTGAACACAGCGAG 1999
QY 1975 GAGTACTATAAGACTGTAAAGACTCAGGACACAGCTCAGAAATACGTCGATATTGAC 2034
DB 2000 GAGTATTACAGAGTGTAAACCTAGTGAACATGAAGGTAGTACTTATATGACATTGAC 2059
QY 2035 GCACGAAAGTGTATTAGCGAAGAACGCAAGTCCCTGTGCTTAAACCGGTATCTGTTA 2094
DB 2060 AAGAAACATGTGTGAAGAAAGAACTGTGTTACGGGACTAGGTTGACAGAGAACTTGTTC 2119
QY 2095 GATCCACATTTTCAGAGATTTTCCGTCAGAGACTCAGACACGACCAAGCACTCTCAC 2154
DB 2120 GATCCCTCTTCATGAATTTGCAATATGAAGCTTGAAGACACGTCGCGCCGACCTTAT 2179
QY 2155 AAAGTCCCAACATCGAGTCTATGAGAGTCCAGGTTTCAAGTAAATCTGAATCATCAA 2214
DB 2180 CAAGTACCACTATATAGTGTATGCTTCCGGGTCCGGGAAGTCTGGAATTTATAG 2239
QY 2215 AGCGTGTACTAAGAAAGATCTGGTTGTAGTGTGGAAGAAAGAACTGCGCAGAAATC 2274
DB 2240 AGCGAGTCAAAAGAAAGAACTGGTAGTGTGTAAGAAAGAAAGAACTGCGCCAGATA 2299
QY 2275 ATCAGGATGTAAAGAGATGAGAGTATGATGTGCTGTAGAGTCTGTGATTCAGT 2334
DB 2300 ATTAAGGAGCTCAAGAAATGAAGGCTTAACGTCATGATGCGCGAGCTGACCTCAATG 2359
QY 2335 CTTCTAAATGGGGTTTAAACACCCCGTTTAACTCTGTATCTTGAAGAGCAATTCCTGCG 2394
DB 2360 CTGTTAAATGATGACAGACATCCCGTCCAGACTTTGTATCTTGAAGAGCTTTGATGAC 2419
QY 2395 CATGAGGAGCTGTGCTGCACTGATTCGATTCGTAACCTAAGAAAGTGGTATTTGTC 2454
DB 2420 CACGCGGCACTCTCAGGGCTTATATAGCAATTAATGACCCAAAGAAAGAGTGTATGT 2479
QY 2455 GGGACCCAAACAAATGGGCTTCTTAAACATGATGTGCTGAAAGTCAATTTAACAT 2514
DB 2480 GGTGACCCAAAGCAATGTGGCTTCTTCAACATGATGTGCTGAAAGTGCACCTTCAACAT 2539
QY 2515 GACATATGACATGAAGTGTACATTAAGAAAGCATCTCTGAGAGGTGCACAGACTGTAC 2574
DB 2540 GAATATGACATCAAGTTTTCACAAAGCAATATCTCGAGGTGTACCAAGTCAAGTACG 2599
QY 2575 GCAATGCTCTCAGGCTCTTCTACGACCAAGGAAATGAAGCGTTAAACCAATGTGCTAT 2634
DB 2600 TCGGATGTGTACACAGTCTTTTATGACAAAGAAATGAAATTAATCAACCAAGGATTC 2659
QY 2635 AAAATCATATGATTAACACAGGACCAACAAAGCCGACAAAGATGTGATTTAAAC 2694
DB 2660 AAAATCGAAATTTGACAAACGAAAGTACTTAAACAAAGAAAGTGAATTTATCTCA 2719
QY 2695 TGTTCAGAGATGGGTGAACAGCTACAGATTGCTCAAAATACAGAAATCATGAGT 2754
DB 2720 TGTTTTAGGGATTTAGGTTTACCACTGCAATTAATCTCAAAAGAAATGAATTAATGACC 2779
QY 2755 GCGGCTGATGCAAGACTTACCGGAAAGCGTTTATGCTGTCAAGTCAAAAGTCAAC 2814
DB 2780 GCGGCTGCTCAGAGGATGACCGGAAAGTGTGTATGCAAGTTAGTTAAAC 2839
QY 2815 GAGATCACTTACTGCGACACTTCTGAGACAGTGAACGTGTTACTTACGACGACGAA 2874
DB 2840 GAGAACCATTTGTAGCAGACCCACCTCAGAGATGTAATGTGCTGTGACCCGAGCGAAA 2899
QY 2875 AAACGACTTGTGTGAAGAGCTAGCTGTGATCTCCGTGATTAAGAACACTTACAGCTTAA 2934


```
Db 2900 GACAAAGTTGTGTGAAGACTCTTGACAGGGGACCGGTGATTAAGACCTTGACCGCGAG 2959
Qy 2935 TATCCCGGGGATTTTCAGGGCTTCATTGACATCTGGACCGCGGAACAGACGCAATTATG 2994
Db 2960 TACCCCGGAAGATTTCACCTGCAACATGGAAATGGCGGCAAGCAATGATGCATCATG 3019
Qy 2995 GCACGCGTCTTGATTAAGCCGACAGCATGATGTTCCAGAAATTAAGGTCAAGCTGCG 3054
Db 3020 AGACACATCTGGAGAAACCGGATCCACGGAATGTTCTTCCAAAATTAAGCTAATGTTTGC 3079
Qy 3055 TGGGCGAAGGCTTTAGAGCCAGTCTTGCCACGCGCAATTTGCTGACAGACAGAG 3114
Db 3080 TGGGCAAGGCACTTGATCTGTGCTTAAGACAGCGGGATGATTAAGACACAGAGAG 3139
Qy 3115 TGGGAGACCTTGACCCATTTAAGCATGACAGAGCTACTCACTGAATTAAGCTGAC 3174
Db 3140 TGGAAACAGTGTATTTCTCAAGAGATTAAGGCCCTCAGCTGAGATTTGCTGAT 3159
Qy 3175 TTTCTTTGACACAGGTTCTTTGAGTAGACCTGACAGTGGGTTATTTTCCGCTCTAC 3234
Db 3200 CAGCTGTGCGTGCATTTCTTGCTGATGACTTAATTTGTTGTTTTCGCCCCACA 3259
Qy 3235 GTCCGACTTACTTACAGGATCAGCACTGGGATTACTGCGGAGAAAGATGATGCG 3294
Db 3260 GTTCACCTCTCATTAAGAACACCATTTGGGACAACTCAGCCTACCCCATGATGCGG 3319
Qy 3295 CTTAATAGAGAGTGAAGAGAGATTGTCAAGCGGATATCCGTGATCAAAAGCGTT 3354
Db 3320 TTGAATCAAGAGTGTGTGAGCAACTATCAGCAGGTAACTCTCAATTAACAGTGCCTG 3379
Qy 3355 GACACAGGCGGATAGTGTATTAAGGAATTAATCAATCAAGAGCTACTCCAACAATT 3414
Db 3380 ACTACTGGAGAGATACGATGAAACCGGTACTTTGGCAATTATGATCCGCGATA 3439
Qy 3415 AATGTGTTCCATTAAATCCGCGGTGCGCCACTGTTGATCGTTGACCAAAAGAGAG 3474
Db 3440 AATTAGTACCGGTGAACCGTGTCTACCATGCTGTGTGAGCAATGCTGATCAT 3499
Qy 3475 GGTACAACTGATCAGCGGATTCCTATCTTAAGATGAAGGCAAACTGTGTGTGATC 3534
Db 3500 CCTCCAGTGAATTTTCGCTTTGTCAGTAAAGCTTAAGGCAAAAGCTCTGTGATGTT 3559
Qy 3535 GGCATCTTATCAGCATTTCCAGGAGAAAGTAGAGTCATGCGGCTCCCACTAAT 3594
Db 3560 GGTAGAAGATGATTAATTCAGTAAGCGTAACTGTGTTATCTGAACACCTGATTTCT 3619
Qy 3595 ACCATCAGGTGATCTCGATTTGGGAATACCTAGCCATGTGCTAAATATGACATTATC 3654
Db 3620 ACTTTAGGGGCGCCTAGATCTAGGCATACCCAATGAATACGAAAGTACGATATGCTC 3679
Qy 3655 TTTGTCAATGTTAGACCCCGTACAGAAACATCACTACCAAGTGGAGATCAAGCT 3714
Db 3680 TTCTTAATGTAAACACAGTACCGCTACCACTACAGCAGTGTGAGAGACCAAGCC 3739
Qy 3715 ATCCACACAGCATGCTAAGGCTGTCCACCACTGAACACTGAGGAGAAACATGT 3774
Db 3740 ATTAAAGTGAACATGTTGACCAAGAGGCTGCTGCACTGAACCCCGAGAACTGT 3799
Qy 3775 GTGGCTATAGGATAAGGCTTTGCTGATGCGCAACCGAATATCATCACTGCGGTGCA 3834
Db 3800 GTGAGCATTTGGTTAGGCTATGCGGACCGGCACTGAACATCATAGTGTGAGTTGCT 3859
Qy 3835 CGCTCATTTAGGTTTACCCGTGTCTGTACGCTTAAGAACTGCGGAAATATGAGGTT 3894
Db 3860 CGGAGATTCAGATCTCGAGGGTATGCAACCGAAGGTGTCTAAGAGAGAACCGAAGTG 3919
Qy 3895 CTCTTCGTTCTTCCGCAAGAGCAACGCGCAACCACTACATGACAGAGAGATCGGCT 3954
Db 3920 CTATTTGCTTCAATGGGTTGATGCTAAACGGAACCAATACCATACAGCTCTCC 3979
Qy 3955 GTAGTGTGCAACATCTATCAAGGTCACAGGTACGAGGAGGAGAGCTCCAGCG 4014
Db 3980 TCCACCTGACCAATATTTACACCGGCTCGAGGCTCCATGAAGCTGCTGCGACTTGC 4039
Qy 4015 TACAGATGATCAAGGTTGATTAAGCAAGAGGCTGACCAAGCTATCTTAATGCTGT 4074
Db 4040 TATCATTAAGTGGCGGGGATATAGCTCACTGCCACGGAAGAGTAAATGTTATCTTCC 4099
Qy 4075 AATGCAAGAGTCAACAGGTTCCGAGTGTGCGTGCACCTGATACGMAAATGCGCGCT 4134
Db 4100 AACGCAAGGCGCAGCCAGGCAATGAGTGTGGGAGCTGTATACGGAAATACCCGAA 4159
Qy 4135 GCTTTGATAGACCCCAATAGCTGTGTGGACGCTGACTGTTGAA--GCACGAACG 4191
Db 4160 AGCTTCATTTTACCAACCAATGAAGTGGGAAAGCTGATTTGTCTCAAGGTAATCTCAAA 4219
Qy 4192 CTATCATATGCTGTAGAGACCAATTTTCTTAAGATGCGCGAACCAGGAGGACCTT 4251
Db 4220 CATCTCATTAAGAGTGGGCGGAATTTTAACAAAGTGTGAAAGTGAAGTGAACAA 4279
Qy 4252 AAGCTGCACTGCTCATGATGATAGCTAGCTTCATCTCAACGCTGACGGAATTACAAA 4311
Db 4280 CAGCTGCAAGAGTATGAAATCTATGACAGGATTTAATGACAAATTAATGATCT 4339
Qy 4312 ATATCAGTACCGTACTGTCAACCGCATCTATTGTGTGCAAAAGATCAAGTATGCA 4371
Db 4340 GTGGCTATTCGGCTTCTGTCCACTGGAATTTTGGCCGAAACAGGATGATTAATCAA 4399
Qy 4372 TCATTCATACCTGTTCACTGCTTTCGACACTACGGAATCCGATGTCACATTAATGCG 4431
Db 4400 TCTTAACCATCTGTTAAAGGCAATGACACAAACGAGTGTGCCATATCTGC 4459
Qy 4432 TTGATTAACAAATGAGAACAGGATTAATCAGGCTTACCGGCAAAAGAGCTGCA- 4490
Db 4460 AGAGCAAGAAATGAGGAAGTGAAGCTTGAAGAGGTGTAGCAGAGAGAGGCGGTAG 4519
Qy 4491 -----AATTGATGATGACAAAGCCAGTACATTAATCT-----TGTGAGGCTC 4536
Db 4520 GAGATATGATCTCCGAAGATCTCCGTAGCAGAGCCGATGAGAGCTGTGTTAGATT 4579
Qy 4537 CACCAAGAGCTTTTGGAGGAGACGAGGTACTCCGTCAATGAGGCAAGTTGAT 4596
Db 4580 CACCTTAAGAGCTTTGGCTGGAAGAAAGTTACAGCACTAGGATGGAAGACATTC 4639
Qy 4597 TCATACCTGAAGGTATACCATTCATCAGACCGCCAGAGCAATTCGCAATTCATGCA 4656
Db 4640 TCATATTTGAAGAACCAATTTTCATCAGGCGCGGAAGACATGCGAAATTAAGCT 4699
Qy 4657 ATGTGCCCAACCAATGTGAGGCTAATGACAGATTTGCTGTGATCTCTGGGGAGAGT 4716
Db 4700 ATGTGGCTGCGCTACAGAGGCTAATGAGCAGGTGTCTTATACATTTGCGGTGAAGT 4759
Qy 4717 ATGTCCAGCATCCGCTCCAAATGCCAGTAGAGAGTCAAGGGGTGCTGCCACCTCAC 4776
Db 4760 ATGAGCAGTAAAGATCCAAATGCCCGTTGAGAGTCAAGAGGATCACCCACCAAGT 4819
Qy 4777 ACACTTCATGCTGTTGATTAATTAAGCTATGACGCTAGGCGGTATCAGGTTGCGCTCT 4836
Db 4820 ACATGCTTGTGTTGTGATCAACGCTATGACCCCGGAAGGGTTCAGGCTTTGAAGCC 4879
Qy 4837 GCGAAGAAAGACATTCGCGGTATGCTCATCTCTGTTGCCGAAGTACAGATAC 4896
Db 4880 TCCCGCCCCCAAAATTAAGTGTGTTCTTCTCCATTTGCGGAAGTACAGAAATACA 4939
Qy 4897 GCGGTGAGAGCTACAGTGCAGAAACCAAGTCTGTTTTCAGGCGTACCAACCGGCT 4956
Db 4940 GAGGTGAGAGATTCATATTTGCAATCTTCTCTCTTAAGTACCTGAGTAC 4999
Qy 4957 GTACACCCAGAGATAC 4974
Db 5000 ATACACCTTAAGAAATAC 5017
RESULT 8
AAx78130
```


1405 GGCTTCCGCGTAGGCTCAGAGTGTGCTGTTGAACCACTGTCAAACCCGACCGGCTTT 1464
1476 CCACTCAGATACACGATTAAGATGCTTTTGGCCAAAGAACCAAGC---GAGATTAA 1532
1465 ACAATGGCCGATGTGAGCATCTGTGCTTACAGCAAGAAAGTGAAAGAGTGCTCA 1524
1533 CCTGTCTCGACGGGTGTGAGCCAGGAGTGCTGAACAAGAGGAAAGAGGTGGAG 1592
1525 GCGGAAGAGATCAGAGAAAGCCCTGCAACCTTGTCTCCCTGAAATAGAAAAAGACCTA 1584
1593 GCCAGCTGACTAGAGAAAGCCCTTACCAACCTGTCTCCATCGCGCCGCGAGACGGGA 1652
1585 GAGGAGAAAGTGAAGCTATTATGCAAGAG-----GAGAGAGAGTGAAGGTGGAGACA 1638
1653 GTCTCGACGTGAGCTTGAAGAACTAGAGATACAGAGAGTGACGGGTGTGGAACA 1712
1639 CCACGAGAGACATCAGGGGTGACAAAGTTACCCAGCGAAGAGAAAGATTGGTCTTACGCT 1698
1713 CCTCGACGCGGTTGAAGATCACCGCACAGCGAAGCATACTAGGAAATTTAGCTA 1772
1699 ATACTTTACCCCGAGCGGTATTGAATGTGAAAACTGGCGTGTATCCACCAATTGGCG 1758
1773 GTTGTCTCCCGCAGACCGTGTCTCAAGAGCTTCAAGTTGGCCCCCGTGCACCTCTAGCA 1832
1759 GAACAAAGTACTGTGAATGACTCACAAAGGTAGGGCAGGAGATACAAAGTGCAGCATAC 1818
1833 GAGACAGTGAATAATTAACATACACGAGAGGCGCGGCTTACAGGTGACAGATAT 1892
1819 CAGCGTAAGTCTATTGTAACGAAGAGCAGCGGCTCCCTGTTCAAGACTTCAGCATTTG 1878
1893 GACCGCAGGGGTCTACTACATGTGATCGGCATTCGGTCCGTGAGATTTCAAGCTTTG 1952
1879 AGTGAAGCGCTACGATGTTTCAACGAGAGGAGTTCGTAACAGATACCTGCACTAC 1938
1953 AGCAGAGCGCGCTATGTGTACAAACGAAGGGGTGTCAACAGGAAACTATACAT 2012
1939 ATCCGAATCAACGGAGGAGCGCTAAACACTGACGAGAGTACTTAAAGCTGTAAAGCT 1998
2013 ATTGCGCTTACGAGACCGTCCCTGAACACGACGAGAGAACTACGAGAAATGACAGCT 2072
1999 CAGAGACAGACTAGATAAGTCTTGTGATTTGACGACGAAGAGTGTGTTAGCGAGAA 2058
2073 GAAGAAGTGAAGCGAGATAAGTGTGACGTAAGTAAGAAATGCTGTGCAAGAGAG 2132
2059 GACCGAGTCCCTGTGTGCTTAAACGGGTATCTGTGATTCACCATTTACAGATTTGCG 2118
2133 GAACCGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2192
2119 TACGAGAGTCTCAGACACGACGACGACCTTCACAAAGTCCCAACCATCGAGTCTAT 2178
2193 TACGAGAGGCTGAAGATACAGCGCTCGGACACATTAAGACTACAGTATGAGAGTCTTT 2252
2179 GAGATGCCAGGTTCAGGTAATCTGGAATTCATCAAAAGCGGTGTGACTAGAAAGATCTG 2238
2253 GGGGTTCGGGATCAGGCAAGTGTGTATTAAGAGCTCTGTGACCAAAACGATCTG 2312
2239 GTTGTGAGTGCAGAGAGAGAACTGCGCAGAAATCATCAGGAGTGAAGAGATGAGA 2298
2313 GTCAACGCGGCAAGAGAGAACTGCCAGAAATAGTTAAAGAGCTGTAAAGAACACCGC 2372
2299 CGTATGAGTGTGTGCTAGAGACTGTGATTCAGTGTCTTAAATGGGTTAAAGACACCC 2358
2373 GGGAAAGGAGACAGTAAAGGAAAACAGTATCTCATCTGTCTAAACGGGTGTCTGTGCGC 2432
2359 GTTAAACTCTGTATCATTTGATGAGCATTTGCTGCCATGACGAGGACGCTGTGCACTG 2418
2433 GTGACATCTTAATGTGAGAGAGGCTTTGCTTGCATTCGCGATCTCTGCGCCCTTA 2492
2419 ATTGCACTGTCAAACT---AAGAAAGTGTATTTGTGCGGAGACCCAAACATGCGGC 2475
2493 ATTGCTCTTTTAAACCTCGAGCAAGTGTGTATGTGCGAGACCCCAACAAATGCGGA 2552

2476 TTCTTTAACATGATGTGCTGAAGATGACATTTTAAACCATGACATATGACTGAAGTATC 2535
2553 TTCTTCAATATGATGACGCTTTAAAGTGAATCTTACACCAACAACTGTGACATGAATGT 2612
2536 CATTAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATGTCTCCAGCTCTTC 2595
2613 CATTAAGTATATCAGACGCTTGACCGCTCAAGTCAAGCCGATGTGTCTACGTTCCAC 2672
2596 TACGACAGGGAATGAAGAGGTTAAACCAATGTGTGATTAATTCATATGATACACA 2655
2673 TACGAGAGCAAGTGCAGACACCAACCGTGTGACAAACCATATCATAGACACACA 2732
2656 GGGACCAAAAGCGGCAAAAGATGATCTGATTTCAACCTGTTTCAGAGAGATGGGTAAA 2715
2733 GAGACAGCAAGCCCAAGCCAGAGACATGTGTATACATGCTTCCAGAGCTGGGCAAG 2792
2716 CAGCTACAGATTGACTACAAAAATCAGAAATCATGCTCGGCTGTGATGCAAGACTT 2775
2793 CAGCTCAGTTGACCTACCGTGTGACAGAAATCATGACACAGCATCTCAGGGCTTC 2852
2776 ACCGGAAAGCGCTTTATGTGTGACGTAAGAGTCAAGAGATCACTCTCTGCGAG 2835
2853 ACCGCAAGGGGTATACGCTGTAGCAGAGAGTGAATGAATCCCTGTATGCCCT 2912
2836 ACTTCTGACGACGTGAACGTGTTACTTACACGACAGAAAAAGCATTTGTCTGGAGACG 2895
2913 GCGTGGAGACATGGAATGTACTGTGTGACGCGCATGTAGGATATGGCTGTGTGGAACG 2972
2896 CTAGCTGTGATCCCTGGAATAAAGACATTAACGCTTAATATCCGGGGATTCAAGCT 2955
2973 CTGGCCCGGCATCTCTGATTTAAGTCTCTATCAAACTTCACAGGGTTACTTTACGCGC 3032
2956 TCAATTGAGACATGGGCGGCGGAAACAGAGCCATTATGGCAGCGCTTGTGTAAGCGG 3015
3033 ACATTGGAAGATGGCAAGAAACAGACAAATTAATGAAGTATGAAGACCGGCT 3092
3016 CAGACAGCTGATGTGTTCCAGAAATGAAGTGAAGCTGTGCGGCGAAGGCTTTAGACCA 3075
3093 GCGCTGTGAGACCGTTCCAGAAACAAAGCAGAGTGTGTGGGAAAGACCTGTGTCT 3152
3076 GTCTTGGCAACGGCCAACTTTGTGTGACGAGACAGAGTGGAGAC--GTTGCACCCA 3132
3153 GTCTTGACACTGCGGAAATCAGATTGACAGAGAGAGTGTGAGACATATATTAACGCA 3212
3133 TTCAAGCATACAGAGCTACTGACCTGAATGGCACTGAACCTTTTGTGACAGGTTG 3192
3213 TTTAAGGAGACAGAGTCTTCAAGTGTGTGCTGTGATGAATTAATTTGACCAAGTATC 3272
3193 TTGGAAGTACCTGACAGTGTGATTTTCCGCTCTTACCGTGTGACTTACTTACAGG 3252
3273 TATGGAATGACCTGTGACAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3332
3253 GATCAGACTGGGATTAATCTGCCAGGAGAAACATGTATGTGGCTTAATAGAGAGTGA 3312
3333 AACCAACACTGGGATTAACAGACTGTGTGAAGATGTATGATTCATATGCGGAAACGCT 3392
3313 AAGAGTGTGACAGGCGATATCGGTGATCACAAAAGCGGTGACACAGGAGGTAGCT 3372
3393 GCCAGCTGAGAGCTACATACCTTCTGAAAGGCGAGTGGCAATAGGGCAAGCAGGCA 3452
3373 GATTAAGAAATATACATCAAGAGACTACTCTCCAACTTAATGTGTGTTCAATTAAT 3432
3453 GTTATGCGAAGAAAGAAATCCAAACCGCTTTGTGTGTGACATATTAATTCATATCAC 3512
3433 GCGCGGTGCCCCACTCTGTTGATGTTGACCAAAAGAGACAGGTTCAATGATCACAGC 3492
3513 CGCAGGCTGCGGACCGCTGTGTGTGAGTGAATGAAGCGTTAAAGCAATAGGTTGAG 3572
3493 GGAATCTATCTTAAGATGAAGGCGAAATCTGTGTGTGATGCGGCACTCTATCAGACTT 3552
3573 TGTGTGTCAATTAAGTAAGAGGTATCAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 3632
3553 CAGGGAAGAAATAGAGTCCATGGGTCCATTGCCCCCACTAATAC-----CATCAGG 3603

Db 3633 TTGCTCGACGAGGATCACTTGTGTGTACCGCTGAATGTACAGCGCCGATAGTGC 3692
Qy 3604 TGTGATCTGATTTGGAAATACCTAGCCATGTCCGTAATATGATATATCTTTGCAAT 3663
Db 3693 TACGACCTTAAGTTTAGACTGTCCGGCTGACGCCGGCAGGTGTGACTTGTCTTTGTGAAC 3752
Qy 3664 GTTAGAACCCCTAGAGAAACCATCATCTAACAGTGGAGATATCCGCTATCCACAC 3723
Db 3753 ATTCAACGGAAATTCAGAAATCCACCATCTACAGAGTGTGTGACCAAGCCATGAAAGCTG 3812
Qy 3724 AGCATGCTTAAGCTGTAAAGCTGTCCACCACTGGAACACTGGCGAAACATGTGTGCTATA 3783
Db 3813 CAGATGCTTGGGGAGATGCGCTACGACTGTAAACCCGGCGGATCT---TGAATGAGA 3869
Qy 3784 GGGATGGGCTTGTGATGGGCAACCGGAATATCATCATCTGCGGTGCACGCTCATTT 3843
Db 3870 GCTTACGGATACCGGATTAATAATCAGCCGAAGCCGTTGTCTCTTAAGCAAAAGTTC 3929
Qy 3844 AGGTTTACCCCTGTGTCTGACGCTTAAGAACACTGGCGAAATATGAGGTTCTTTGCTG 3903
Db 3930 TCGTCTGCAAGAGTGTGGCCCGGATTTGTCTACCAAGAAATACAGAAAGTCTTTGCTG 3989
Qy 3904 TTCTTGGGAGAGCAACGCGCAACACATGACAGAGACACTGCGTGTAGTCTT 3963
Db 3990 TTCTCAACTTTGACAAACGAAAGAACCTCTACGCTACACAGATGAATACCAAGCTG 4049
Qy 3964 GACAACTCTATCAAGGCTCAACAGGTACAGGTACAGGAGGAGGAGCTCCAGCTTACAGAGTG 4023
Db 4050 AGTGCCGTGTATCCGGAGAAACCATGACACGCGCGGTGTGTACCATCTTACAGAGT 4109
Qy 4024 ATCAGAGGTGACATTAGCAAGAGCCCTGACCAAGCTATGCTTAATGCTGTATATGCAAA 4083
Db 4110 AAGAAAGGAGATATAGCCACGTGTGCAAGAAAGGCTGTGTATAGCGACTAACCCCT 4169
Qy 4084 GGTCAACAGAGTTCCGAGTGTGCGGTGCACTGTACCGAAATGCGCGCTCTTTGAT 4143
Db 4170 GGAACCTGTAGGGGATGCGGTATGCAAGGCGCGTGGCAAGAAATGCGCCGACCTTTAAG 4229
Qy 4144 AGACAGCCCAATGCTGTGGGAGCGGCTAGACTTGTGAAGACGAAACG---CTCATCATTA 4200
Db 4230 GAGACAGCAACACAGTGGGCAAAATTAACAGTATGTGTGGCTGTGATCCCGCTCATC 4289
Qy 4201 CATGCTTAGAGACCCCAATTTTCTAAGATGCCGGAACCGAGGCGACCTTAAGCTGCA 4260
Db 4290 CACGCTGTAGGCGCTTAATTTCTGTGCAAGATGTAAGCGGAAGCGCAATTTGGCC 4349
Qy 4261 GCTGCTTACATGAGATAGCGCTCATTCGTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db 4350 GCTGCTACCGGGCAGTGGCCGCCGAAGTAAACAGACTGTCACTGAGCAGCTAGCATC 4409
Qy 4321 CCGCTACGTGTCAACCGGATCTATTTGTGTGGCAAAATGATGATGATCATTTGAT 4380
Db 4410 CCGCTGTGTGTCAACGAGGTGTTCAGCGCGGAAAGATAGGCTCAGCAATCCCTCAAC 4469
Qy 4381 CACCTGTTCACTGCTTTGAGCACTACGGAATGCGATGTCACCATATATTTGCTGATTA 4440
Db 4470 CATCTATTCACGCAATGAGCCCGAGCGGTGACGTGATCCTACTGCAAGAACAA 4529
Qy 4441 CAATGGAGACCAAGATATTCAGAGCCATTCACCGCAAAAGAAAGCGTCAAAATTTGAT 4500
Db 4530 AGTTGGAGAAAGAAATCCAGAAACCATGTACATGAGGAGCGGTGTGATGTTCAAT 4589
Qy 4501 GATGCAAGCCAGTATGACATTAATTTGTGAGGGTTCACCCCAAAACAGCTTTTGGCAGGC 4560
Db 4590 GATGACGTGAGACTACCAACAGACTTGTGAAAGTGCACCCGACAGCAGCCCTGTGGGT 4649
Qy 4561 AGACAGGTTACTCCGTCAATGAGGCAAGTTTATTAATCTGGAAGGTACAGATTTC 4620
Db 4650 CGTAAGGGCTACAGTACATGACGAGGTGCTGTACTGTACTTTGAAGATTCGAAATTC 4709
Qy 4621 CATGACCGCCAGAGCAATTCGCAAAATTCATGCAATGTGGCCCAACAAATCTGAGCT 4680

Db 4710 AACGAGCTGCTATTGATATGACAGATACTGACGTTGTGGCCAGACATCGCAAGAGCA 4769
Qy 4661 AATGACGATTTGCTGTATACATCTGGGGGAGATATGTCAGCATCGCTCCAAATGC 4740
Db 4770 AACGAAAGATATGCTTATAGCGCTGGCGAAACATGACATCATCATCAATGT 4829
Qy 4741 CCAGTAGAGAGTCAAGAGCGCTGTGCTTCCACTCAACACTTCCATGCTGTATATAC 4800
Db 4830 CCGGTGAACGATTCGATTCATCAACACTCCAGAGCAGTGCCTGTGTCGGCTAC 4889
Qy 4801 GCTATGAGCGCTGACCGCTTATCAGGTTGCGCTTGCAGAAAGAAACAGTCCCGTA 4860
Db 4890 GCAATGACAGCAAGACGATGCGCCCTTGTGTACACCAAGTTAAAGCATGTGTGTT 4949
Qy 4861 TGTGATCATTCCTGTTCCGGAATGACAGATCAACAGCGGTGACAGACTACGTGACGC 4920
Db 4950 TGTGATCTTTTCCCTCCGAAATACATGATGATGGGGGTGACAAAGTAAAGTCCGAG 5009
Qy 4921 AACCAAGTCCGTTTTCAGGCGGTGTACACCGGCTGTACACCCGAGAAATGACGAG 4980
Db 5010 AAGGTTCTCTGTTGACCGCAGCGTACCTTCAAGTGTATGTCGGCGGAAGATGCCCA 5069

RESULT 9
AAC55473
ID AAC55473 standard; DNA; 12464 BP.
XX
AC AAC55473;
XX
DT 11-JAN-2001 (first entry)
XX
DE Destination vector pDEST9 nucleotide sequence.
XX
KM Bacteriophage lambda; att; recombination site; attB; attP; attR; attL;
KM mutant; recombinational cloning; entry vector; destination vector;
KM gene product targeting; fusion tag cleavage; ds.
XX
OS Bacteriophage lambda.
OS Synthetic.
XX
PN WO20052027-A1.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-US05432.
XX
XX 02-MAR-1999; 99US-0122389.
PR 23-MAR-1999; 99US-0126049.
PR 28-MAY-1999; 99US-0136744.
XX
PA (LIFE-) LIFE TECHNOLOGIES INC.
XX
PI Hatley JL, Braach MA, Temple GF, Cheo D;
XX
DR WPI; 2000-543948/49.
XX
PT Isolated nucleic acid molecules encoding an attB1, attB2, attP1, attP2,
PT attL1, attL2, attR1, and attR2 nucleotide sequence useful for the
PT recombinational cloning of polypeptides -
XX
PS Disclosure; Fig 29; 459pp; English.
XX
CC The present invention describes isolated nucleic acid molecules (I)
CC encoding an attB1, attB2, attP1, attP2, attL1, attL2, attR1, and attR2
CC nucleotide sequence. Also described are: (1) an isolated nucleic acid
CC molecule (II) comprising one or more att recombination sites comprising
CC at least one mutation in its core region that increases the specificity
CC of interaction between the recombination site and a second att
CC recombination site; and (2) an isolated nucleic acid molecule (III)
CC comprising one or more mutated att recombination sites comprising at
CC least one mutation in its core region that enhances the efficiency of
CC recombination between a first nucleic acid molecule comprising the
CC mutated att recombination site and a second nucleic acid molecule

Db 6141 GATTCCACCTGAAGGTAACAATCCTTACCTGTAGGTGGGATACCATCGTATCATGT 6200

Db 7218 AGCGAGAGCGCCACTATGTTGTACACGAAAGGAGTTCGTCAACAGGAACATATACCAT 727

QY	1939	TTCCGAAATCAACGGAGGAGCCCTAAACCTGACGAAAGGTACTATTAAGCTGTAAAGCT	1998
Db	7278	ATTGCGGTTCAACGGACCCTCGGTGAACCCGACGGGGAATCTAGAAAATGCAGAGCT	7337
QY	1999	CAGGACACAGACTCGAATTAAGTCTTCCGATATGACGACGGAAGGTGTTAAGCGAAG	2058
Db	7338	GAAAGAACTGACGCCGAGTACGTGTTCCAGCTAGATTAATAAAATGCTGCTCAAGGAAG	7397
QY	2059	GACGCAAGTCCCTTGTGCTAAACCGGTATCTGTAGATCCAGCAATTTTCAGAGTTTGG	2118
Db	7398	GAAAGGTGGGTTGTGTGTGGTGGAGAGCTAACCAACCCCGCTTCATGAATTCGCC	7457
QY	2119	TACCGAAGTCTCAGAACACGACGACGACACTCAAAATGCCAACCATGTGGAGTCTAT	2178
Db	7458	TACGAAAGGCTGAAGAAATCAGGCGCGTCCGACCATTAAGACTACGTAGTAGAGTCTTT	7517
QY	2179	GGAGGCGCAGGTTCAAGTAAATCTGGATTCATCAAAAGCGGTGACTTAAGAAAGATG	2238
Db	7518	GGGGTCCGGGATCAGGCAAGTCTCTCTATTATTAGAGCTCTGTGACCAACAGATGTG	7577
QY	2239	GTTGTAGTGCAGAAAGAAAACTCCGACAGAAATCATCAGGGAATGAAGAGATGAGA	2298
Db	7578	GTCACACACGGGCAAGAAAGAGAACTGCCAGGAAATAGTTAAAGACTGAAGAACGCC	7637
QY	2299	CGTATGATGTTTGTCTGCTAAGACTGTGCATTCAGTCTTCTTAATGGGGTTAAGCACC	2358
Db	7638	GGGAAAGGGGCAAGTAGAGGAAAAACGTACTCCATCTGCTTAACGGGTGTCTGTGTC	7697
QY	2359	GTTAAACACTCTGTACATTGATGAGGCAATTCGCTGCCAATGACGAGACGCTGTGCACTG	2418
Db	7698	GTGGACATCTTATATGTGTGAAGAGGCTTTCGTTGCCATTCGCGTACTCTGTGCTCTA	7757
QY	2419	ATTGCGCATCTCAAACT---AAGAAATGGTATATGTTCGGGGGACCCAAAACAAATGCGGC	2475
Db	7758	ATTGCTCTGTTAAACCTCGGAGCAAAATGTGTATTATCGGAGACCCCAACAAATGCGGA	7817
QY	2476	TTCTTTAATCATGATGTGCTTGAAGTACATTTTAAACATGACATATGCACTGAAGGTAC	2535
Db	7818	TTCTTCATATGATGACAGCTTAAGGTGAATTTCAACCAATCTGCACTGAAGTATGT	7877
QY	2536	CATAAAGCATCTCTAGAGAGTGCACACAGACTGTAACCGGCATGTGTCCAGCTCTTC	2595
Db	7878	CATAAATATTAATCCAGAGCTTGACGGGTCCAGTCAAGGCGCATGTGTCTACGTTGCAC	7937
QY	2596	TACGACAAGCAATGAATGAACGTTTAACCATATGTGTGATTAATAATCATGATGATACCA	2655
Db	7938	TACGAGGCAAGATGCGACACACCAACCCGTGCACAAACCATATATATGACACACCA	7997
QY	2656	GGGACCAACAAAGCCGACAAAGATGATGTGAATTTCAACCTGTTTCAGAGATGGGTGA	2715
Db	7998	GGACAGAACCAAGCCCAAGCCAGAGACATGCTGTTTAAATCATGCTTCCGAGGCTGGGCAAG	8057
QY	2716	CAGCTACAGATTGACTACAAAAATCAGAAATCATGATCTGGGGCTGTGTCAGCAAGACTT	2775
Db	8058	CAGCTGCACTTGACTACCGTGCACAGCAATCATGACAGACGACATCTTCAGGCGCTC	8117
QY	2776	ACGCGGAAAGGCGTTATGTCTGTCAAGTACAAAGTCAACGAAATCCACTCTACTCGCAG	2835
Db	8118	ACCGGCAAAAGGGGTATACGCCGTAAAGGAGAAATGTAATAATCCCTTGTATGCCCT	8177
QY	2836	ACTTCTGAGCAGTGAAACGTGTACTTACACGCAAGAAAAACGATGTCTGGAAGACG	2895
Db	8178	GCCTGCGAGCAGTGAATGTACTGTGACGGCACTGAAGATAGGCTGTGTGGAAAAAG	8237
QY	2896	CTAGCTGTGATCCCTGTGATTAAGACACTTACAGTAAATATCCCGGGGATTTCAAGGCT	2955
Db	8238	CTGGCGGCGATCCCTGTGATTAAGGTCTTATCAAACTTCCACAGGGGTAACTTTACGGCC	8297
QY	2956	TCAATTGACGACTGGCAGCGGAACACAGACGCAATTATGGACGCGTTCTTGATTAAGCG	3015
Db	8298	ACATTGGAAATGTGCAAGAAAGACACCAAAATATGAAGGTGATTTGAAGAACCGCT	8357
QY	3016	CAGACAGCTGATGTGTTCCAAATTAAGGTGAACGTCTGCGGCGAAGCTTTAGAGCCA	3075

Db	8358	GGCCCTGTGGACGCGCTTCAGAAACAACGAAACGTGTTTGGCCAAAGCCCTGTGCTT	8412
Qy	3076	GTCTTGGCCACGGGCAACCATTTGTGTGACGAGACAGCAGTGGAGAC---GTGGACCCCA	3132
Db	8418	GTCCTGGACACTGCCGGAAATCAGATTGACAGACAGAGAGTGGAGCACCATAATTAACAGA	8477
Qy	3133	TTCAAGCATGACAGAGCCGTAATCACTTGAAATGGCACTGAACTTCTTTGGACCAAGTTC	3192
Db	8478	TTTAAGAGGAGACAGAGCTTACTCTCCAGTGGTGGCTTGAATGAAATTTGGACCAAGTAC	8533
Qy	3193	TTTGGAGTAGACCTGGACAGTGGGTTATTTTCCGCTCCTACCGTCGACCTTAATTACAG	3253
Db	8538	TATGAGATTGACCTGGACAGTGGCTGTGTTTCTGCCCCGAAGGTGTCCCTGTATTACAG	8597
Qy	3253	GATCAGCACTGGGATTAATCGGCAAGGGAAGAACTGTATGGGCTTAATAGAGAGGTACGA	3312
Db	8588	AACAACCACTGGGATTAACAGACTGGTGGAAAGATGTATGATTTCAATCCGCAACACT	8655
Qy	3313	AAGAGTTGTCAACGGCGATATCCGTGATCAAAAAGCGTTGACACAGCAGGGTAAGT	3372
Db	8658	GCCAGGCTGGAAGCTAGACATACCTTCTGTAAGGGGCGTGGCATATGGGCAAGGACGCA	8711
Qy	3373	GATATAGGAATTAATACATCAAGGACTACTCTCCACAAATTAATGTGCTTCATTAAT	3433
Db	8718	GTTATCGCAAGAAAGAAAATCCAACCGCTTTCTGTGCTGACCAATGTAAATCTATCAAC	8777
Qy	3433	CGCGGGTGGCCCACTCGTATGTGTTGACCAAAAGGACAGGGTACAATGATACAGC	3492
Db	8778	CGCAGGCTGGCCGACGCCCTGTGGCTGATGTAACAAGCGTTAAAGCGTAGGGTTGAG	8837
Qy	3493	GGATTCCTATCTTAAGATGAGTAAAGGGCAAACTGTGTGTGATCGCGCATCTATACAGATT	3552
Db	8838	TGGCTGTGCATTAATGAATGAAGGGTAAACAGGTCCTGCTGTGATGAGTACACCTGGCT	8897
Qy	3553	CCAGGGAAGAAAGTAGAGTCCATGGGTCATTGCCACTAATAC-----CATCAGG	3603
Db	8898	TTGCTTCACACCAGGCTCACTTGGTTGTCAACCGCTGAATGTCAAGCGCCGATAGTGC	8957
Qy	3604	TGTGATCTGCATTTGGGGAATACCTAGCATGTGGTAAATATGACATTAATTTGTCAAT	3665
Db	8958	TACGACCTTAAGTTTGAAGCTCCCGGCTGACCCCGCAGTTTCGATGTGTTGTGTGAAC	9017
Qy	3664	GTTAGGACCCCGTACAGGAACCATCACTACCAAGTSCGAGATTCACGCTATCCACCAC	3723
Db	9018	ATTCAACAGGAATTCAGAAATCCACCCTACCAAGCAGTGTGTGCACACAGCCATGAAGCTG	9077
Qy	3724	AGCATGCTAACGTGTAAAGCTGTCCACCACCTGAACACTGGCGGAACATGTGTGGCTATA	3783
Db	9078	CAGATGCTTGGGGGAGATGGCGCTACAGCTGTAAAAACCCGGCGGATCT---TGATGAGA	9134
Qy	3784	GGGATATGGGCTTGTCTGATCGCGCAACCGAATATCATCACTGGGGTGGACGCTCATTT	3843
Db	9135	GCTTAAGGATACGCGGATTAATACGCAACCCGTGTGTTCTCTTAAGCAGAAAGTTC	9194
Qy	3844	AGGTTTACCCGTGTCTGTCAAGCTTAAGAACTGCGCCGAATAATCTAGAGGTTCTCTTG	3903
Db	9195	TCGTCTGAAGAGTGTTCGCCCGGATTTGTGTACACAGCAATTAAGAAAGTGTCTTGCTG	9254
Qy	3904	TTCTTGGCAAGAACAGGCAACCAACATGACAGACAGACAGACTCGGTGTAGTGTCTT	3963
Db	9255	TTCTTCAACTTTGACCAAGAAAGAGACCTCTACGCTAACAGATGAATACCAAGCTG	9314
Qy	3964	GACAACTCTATCAAGGGTCAACCAAGTACGAGGCAAGGAGAGCTCCAGCTACAGAGTG	4023
Db	9315	AGTGCCTGTATGCCGGAAGGCAATGACACGGCCGGGTGTGACCAATCTTAAGAGATT	9374
Qy	4024	ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCGTTAATGTGTCTAATAGCAAA	4083
Db	9375	AAGAGAGGACATTAAGCAGCTGACAGAAAGCGGCTGTGTAAACGACCTAACGCCCT	9434
Qy	4084	GGTCAACCAAGTTCCGAGTGTGCGGTGCACTGTACCGAAATAGCCGCGCTTTTGAT	4143


```
D 9435 GGAACGTGAGGAGATGGCGTATGACGAGGCGGTGGCGAAGAAATGCGCTTAAAG 9494
Q 4144 AGACAGCAATAGCTGTGCGGACGGCTAGACTTGTGAAGACGACG---CTATCATTA 4200
D 9495 GGACAGCAACACACGAGTGGGACAAATTAAACAGTCAATGTGCGCTGTACCCCGTATC 9554
Q 4201 CATGCTGAGGACCAATTTTCTAAGATGCGGAAACCGAGGGGACCTTAAGCTGCA 4260
D 9555 CACGCTGAGGCTTAATTTCTGTCCACGACTBAGCGAAGGAGCGGGAATGGCC 9614
Q 4261 GCTGCTCACTAGACATAGCTCATGCTCAAGCGCTGAGCGATTACAAAATATCAATA 4320
D 9615 GCTGTACCGGAGTGGCGCGCAAGTAACAGACTGTCACTAGACAGCGTACCATC 9674
Q 4321 CCGCTACTGTCAACCGGCTATTTCTGTGCGCAAAATGCAAGTACATCATTTGAT 4380
D 9675 CCGCTGTCTCAAGAGATGTTGAGCGGCGAAGATGAGTGGCTGACGACATCCCTCAAC 9734
Q 4381 CACCTGTCACTGCTTTGACACTACGATCCGATGTCACCATATATTTGCTGATATA 4440
D 9735 CATCTATTCAAGCAATGAGACGCAACGACGCTGACGTCATCTACTGACAGACAAA 9794
Q 4441 CAATGGAGACACAGATATTCAGAGGCCATTCACCGCAAAAGAACGTTGAAATTCGAT 4500
D 9795 AGTGGGAGAAATTCAGAGAACCATTCATGATGAGAGACGGCTGTGAGTTGCTCAAT 9854
Q 4501 GATACACAGCAGATAGACTTGTGTCAGGGTCCACCCAAACAGCTCTTTGGCAGGC 4560
D 9855 GATACACGTGAGCTGACACACAGCTTGTGAGTGCACCGGACAGACGCTGTGGGT 9914
Q 4561 AGACAGGATTTCTCCGTCAATGAGGCGCAAGTTATTCATCTCTGAAAGTACAGATTC 4620
D 9915 CGTAAAGGCTACATACCTGACGCGGTGCTGTACTGTGAAAGTACCAATTC 9974
Q 4621 CATCAGACCGGCAAGACATTTGCCAAATCCATGTCATGTCGTCAGGCT 4680
D 9975 AACGAGCTGTATTTGATATGAGACAGATACGATGTCGTCGTCGTCAGAGAGCA 10034
Q 4681 AATAGCAGATTTGCTGTATCTCTGGGAGAGATATGTCCAGATTCCTCCAAATGC 4740
D 10035 AACCAAGATATTCCTATAGCGCTGGCGCAAAACATGACCAATCAATGATTCAAATGT 10094
Q 4741 CCAATGAGGAGTACAGAGCGCTGTCTCACTGACACTTCATGCTGTGTAATTC 4800
D 10095 CCGGTGAACGATTCGATTCATCAACACTCCACAGACAGTGCCTGCTGTGCGCTAC 10154
Q 4801 GCTATGACGCGTGAACGCGTATACAGTTGGCTCTGCGAAGAAAGAACATTCGCCGTA 4860
D 10155 GCAATGACAGCAGAAACGATCGCCGCTTAGTTCACACCAAGTTAAAGATGTGGTT 10214
Q 4861 TGCATCATCTCTGCTGTGCGAAGTACAGATTCACAGCGCTGCGAAGCTACATGACG 4920
D 10215 TGCATCATCTTTTCCCTCCCAATATCCATGTAAGTGGGTCGCAAGTAAAGTGGCAG 10274
Q 4921 AAACAGCTCTGTTTTCAGCGGTGTACACCGGCTGTACACCCGAGAAATACGCGAA 4980
D 10275 AAGGTTCTCTGTTGACCGCAGCGTACCTTCAAGTGTGTTGCTCGCGAAGTATGCCGA 10334

RESULT 10
ACC45139 standard; DNA; 7399 BP.
ID ACC45139;
AC ACC45139;
DT 11-JUN-2003 (first entry)
DE Alphavirus replication unit NS1-NS4 gene related sequence.
XX vaccine; alpha virus replication unit; NS1-NS4; gene; ds.
XX OS Alphavirus.
XX
```

```
PN CN1367020-A.
XX
PD 04-SEP-2002.
XX
PF 19-OCT-2001; 2001CN-0128979.
XX
PR 19-OCT-2001; 2001CN-0128979.
XX
PA (VUHU-) UNIV HUIXI HOSPITAL SICHUAN.
XX
PI Zhao L, Qin S;
XX
DR WPI; 2003-168348/17.
XX
PT Nucleic acid vaccine capable of being high-effectively and safely
PT expressed comprises a conventional nucleic acid vaccine plasmid and
PT alpha virus replication unit NS1-NS4.
XX
PS Example; Fig 1; 14pp; Chinese.
XX
CC The present invention describes a nucleic acid vaccine which can be
CC effectively safely expressed. The nucleic acid vaccine plasmid forming
CC the nucleic acid vaccine is formed from a conventional nucleic acid
CC vaccine plasmid and an alphavirus replication unit NS1-NS4 gene
CC recombination. After having been injected into body, under the action
CC of a CMV promoter, the nucleic acid vaccine can transcribe the NS1-NS4
CC replication unit and RNA of its downstream target gene with high
CC efficiency, and it possesses high expression efficiency of its target
CC antigen, and after its transcription in the cell and expression level
CC has reach a certain high level, it can induce the cell transfected by
CC performing a death process to form a self-destrorying mechanism so as to
CC effectively solve the safety problem of nucleic acid vaccines. The
CC present sequence represents an alphavirus replication unit NS1-NS4 gene
CC related nucleotide sequence, which is used in an example from the
CC present invention.
XX
SQ Sequence 7399 BP; 1966 A; 1915 C; 2025 G; 1493 T; 0 other;

Query Match 31.3%; Score 1564; DB 25; Length 7399;
Best Local Similarity 58.3%; Pred. No. 0;
Matches 2904; Conservative 0; Mismatches 2040; Indels 36; Gaps 8;

Q 31 AGAATTCACGCTTGAATGATGCTGACAGCCCGTATGCTCAAGTGTGTTACACGCGACGTTT 90
D 95 AAGGTGATGTTGATATGATGAGCTGACAGCCCATTCATCAAGTCTTTGCAAGAGCATTT 154
Q 91 CCACAAATTTGATGATGAGCAAGGACGCTACTGACATACATACCAATGCCAATGCCAGACG 150
D 155 CCGTCTGCGAGGTGAGTATGATTCAGAGTCAACCAATATGACATGCCAATATGCCAGACA 214
Q 151 TTTTGGATGCGCAACAAAGCTCATTTGAGAGGAGTGCACGCGGACCAAGTTATCTTG 210
D 215 TTTTGGACCTGCTGATCAATATGATGACAGAGAGACTBACAAAGACACACTCATCTTG 274
Q 211 GACATTTGAAGTGGCCCGCTCAACATGACATTCATTCATACCGCTATCTATCTGTC 270
D 275 GATATGGGAGTGGCTTCTTCCAGAGAAATGATGTCTAGGACAAATACCACTGTGATGC 334
Q 271 CTTATGATTAAGCGCTGAAGACCGGACAGACTACCAACGCTATGACAGAAAGCT-----T 324
D 335 CCTATGGCAGCGCAGAAAGACCCGAAAGGCTCGATGCTACGCAAGAAAGAACTGGCAGCG 394
Q 325 AAGAAAAGTACATTACCGACAGAAACATAGCCCTTAAGCGGAGACCTGTGCGAAGTC 384
D 395 GCCCTCCGGGAAGGTGCTGATGAGAGATGCGAGAAATATCAACCGACTGACAGACGTC 454
Q 385 ATGTCAACACAGACGAGAGACTCATCTGTGTATGACACAGACGACGACGCTGATG 444
D 455 ATGGCTAGCGCAGACGCTGATCTCTTACCTTTTGGCTGATACAGACGCTACGCTGCT 514
Q 445 TACTTTGGAAGTGAAGTATACCAAGATGTGTACGCACTGCATGACGACGACATCATC 504
D 515 ACGGACACCGAAGTGGCGGTATACAGAGACGTGTATGCTGTATCATGACCAACATCCCTG 574
```


505 TACCAACAGGCGCTTAAAGAGTTAGACAATTACTGGATAGGCTTTGACACGACCCCT 564
575 TACCATCAGGGGATGAAAGGTGTGAGAACGGGTAATTGGATTGGTTGACACACCCCG 634
585 TTTATGTACAAAACATGGCAGGCTTCTACCTACTTACACAGCAAGCTGGGCTGACGAG 624
635 TTTATGTTTGAACGGCCTAGCAGGCGCGTATCCAACTTACGCCAACAATGGGCGGACGAG 694
645 AGAGATTGGAAGCAGTACATTTGGCCCTGGTAATCTGAGATCTTCAGAGAGGAGGCTT 684
695 CAGGTGTACAGGCGAGGAGCATAGAGCTGTGTGAGCATCTCTTACCTGAGGAGGAGACTTC 754
685 GAAAACTCTCAATCTTAGAGAAAGAGGCTCAACCTACTAATATAGACTATTTCTCG 744
755 GGGAAACGTCTCATTTCTCCGCAAGAGCAATTGAAACCTTGGACACAGTATGTTCTCG 814
745 GTTGTTCAACATCTACACAGAGATGATCACTGTTACGTAAGTGGCATCTTCCAAAC 804
815 GTAGATCTACATTTGTACACTGAGAGAAAGCTACTGAGGAGCTGGCACTTACCTCC 874
805 GTGTTCCACTTGAAGAAAGAGTCTTACACAGTATAGTGGGACCATTTGTACGCTGT 864
875 GTATTCACCTGAAAGGTAAACAATCTTTACCTGTAGGTGATGATCATCATGT 934
865 GAAGGAGTCACTCAAAAAGATTAACATCAGCCAGGACTATAGCGTAAAGTTAGAAC 924
935 GAAGGTACGTATGTAAGAAATCACTATGTCCCGGCTGTACGTTAAACGTTAGAGG 994
925 TTGGGTCACAAATGCATCGCAGAGGTTCTTGAATTTCAAAATGACAGATACGCTGGC 984
995 TACGCGTACGATACGCGGAGGAGTTCTTGTGTCAAGACACAGACACTGTCAA 1054
985 GGGCAGAGGTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1044
1055 GAGAAAGTCTCAATCCCTGTATGACATACGCTCCCTCAACATCTGTGATCAATG 1114
1045 ACAGGATTTCTGGCACTGACGTTAGTGTGATGACGACAAAACCTATTTGGGCTC 1104
1115 ACTGCACTACAGGACCGAGCTCACCAGAGAGGACAGAACTTTGTGTGTGTGTGTGT 1174
1105 AACCAAGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1164
1175 AATCAGAGATAGTTGTGAACGAGAAACACAGGAAACATTAACATATAGAACTAT 1234
1165 CTATTAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1224
1235 CTGCTTCGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1294
1225 GACGACGAGAAAGATAGGAGGTGTGGAGGCACTTACTATGTGGGCTGTGTGTGTGT 1284
1295 GATGTGAAACCTCTGTGGGTGTCCGAGAGAGTACTTACTGTGTGTGTGTGTGTGTGT 1354
1285 TTCAAGACCCAGAAATCATCATCTTCAAGAAAGCTGTGTGTGTGTGTGTGTGTGTGT 1344
1355 TTTAAAGAGAGAGATGACACCATGTACAAAGAAACAGAACCCACAACTATGTAG 1414
1345 GTACCTGCGCTTTGTGATCTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1404
1415 GTGCTTCAGAGTTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1474
1405 GCGTTCCGCGTAGGCTCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1464
1475 CCACTCAGATCACGATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1531
1465 ACAATGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1524
1532 CTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1591
1555 GCGGAGAGATCAGAGAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1584
1592 GCGGAGTCTAGAGAGCCCTTACACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1651

1585 GAGGACGAATAGACCTCATTTATGCAAG-----GCAGAGCAGTAGCGTGAGACA 1638
1652 GTGCTCAGACCTGTGACGTTGAAAGAACTAGAGATATACCGAGGTGCAAGGGGTGTGAAACA 1711
1639 CCAAGAGACATCAGGGTGAACAAGTTTACCAGGCGAAGAGAAATTTGGTCTTTAGCCT 1698
1712 CCGCAGCGCGGTGAAAGTACCGCACAGCCGAACGACGTACTACTAGAAATTAACGTA 1771
1699 ATACTTTTCAACCCGAGCGGTATTTGAATGTGTAAGAAAACTGGGCTGTATCCACCATTTGG 1758
1772 GTTGTGTCCGACAGACCTGTCTCAAGAGCTTCAAAGTTTGGCCCCCGTGTACCTCTTACCA 1831
1759 GAACAAGTACTGTGTAATGACTCACAAGAGTGTAGGCGAGGAGATACAAAGTGTGAGCATAC 1818
1832 GAGCAGGTGAAATATATATACATTAACGAGAGGCGCGGCTTACAGGTGTGACGATAT 1891
1819 CAGGTAAGTCAATTTGACAGAAAGGACGCGGTCCCTGTTCAAAGCTTCCAGGATTTG 1878
1892 GACGCGAGGGTCTACTACATGTGTGATCGGCAATTCGGTCCCTGAGTTTCAAGCTTTG 1951
1879 AGTGAAGCGCTACGATGTTTTCAGAGAGGAGTGTGTAAACATATACGACAC 1938
1952 AGCAGAGCGCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2011
1939 ATGCAATCAAGGAGGAGCGCTAAACATGACGAGAGTACTTAAAGTGTAAAGACT 1998
2012 ATTTGCTTTCAGAGACCTGTCTGAACCGACAGAGAACTACGAAAGTGTGAGCT 2071
1999 CAGGACACAGCTGAGATACGTCTTGTATTTAGACGACGAAAGTGTGTAAAGCAGAA 2058
2072 GAAGAAGTGAACGCGAGTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2131
2059 GACCGAGTCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2118
2132 GAACGCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2191
2119 TACGAGAGTCTCAAGACACGACGAGACCTCAACAAAGTCCCAACCATGAGTGTAT 2178
2192 TACGAGAGGCTGAAGATCAGGCTCGGACCATATATAGATTAACATGTATGAGTCTTT 2251
2179 GAGTGCAGGTTCAAGTAAATCTGAAATCATCAAAAGCGCTGTGTGTGTGTGTGTGTGT 2238
2252 GGGGTCCGGGATCAGGCAAGTGTGTATTTAAAGGCTGTGTGTGTGTGTGTGTGTGTGT 2311
2239 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2298
2312 GTCAACAGCGCAAGAGAGAACTGTCCAGGAAATGTATTAACGAGTGAAGAACACCGC 2371
2299 GGTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2358
2372 GGGAGGAGACAGTATGAGAAACAGTGTCTCATCTGTCTAAACGGGTGTGTGTGTGT 2431
2359 GTTAAACCTCTGTATCATTTGTAGGCAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGT 2418
2432 GTGACATCTATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2491
2419 ATTGCAATGCTCAAACT---AAGAAAGTGTATTTGTGTGTGTGTGTGTGTGTGTGTGT 2475
2492 ATTGCTCTTTGTAAACCTCGAGCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2551
2476 TTTCTTAAACATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2535
2552 TTTCTTAAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2611
2536 CATAAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATGTCTCCACGCTCTTC 2595
2612 CATAAAGTATATCAACAGCTGTGACGCGTCACTACGCGCATGTGTGTGTGTGTGTGTGT 2671
2596 TACGACAAAGCATGAAAGCGGTTAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2655
2672 TACGAGGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2731
2656 GAGACACAAAGCGGACAAAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2715

```
Db 2732 GGAAGACCAAGCCCAAGCCGAGAGACATCGTGTAAATCTCCGAGCGTGGCAAG 2791
Qy 2716 CAGCTACAGATTGACTCAAAAATCAGGAATCTGACTGGGCTGCATCCGAGGACTT 2775
Db 2792 CAGCTGAGTTGGACTACCGTGAACAGAACTCATGACAGAGACATCTCAGGGCTC 2851
Qy 2776 ACCGGAAGAGCGTTTATGCTGTAGGTAAGTAAGTCAACGAAATCACTACTCGAG 2835
Db 2852 ACCCGCAAGGGGATACCGCTGAAGCAGAGGTGAATGAATTCCTTGTATGCCCT 2911
Qy 2836 ACTTGTAGACGTTGAACGTGTACTTAACGCAACAGAAAACGCAATTTCTGGAAG 2895
Db 2912 GCGCTGAGACGCTGAATGTACTGTACGCGCACTAGAGATAGCTGTGTGAAAAAG 2971
Qy 2896 CTAGCTGTATCCCTGATTAAGACACTTACACTTAATTCCTGGGATTTTCAAGCT 2955
Db 2972 CTGGCCGCGATCCCTGATTAAGGTCTTATCAAACTTCAAGGGTAACTTTACGCGC 3031
Qy 2956 TCATTGACGACTGGCAGCGCAACGACCGCATTTATGCAACCGCTTGTATAGCG 3015
Db 3032 ACATTGAAAGATTGGCAAGAAACAGCAAAATTAATGAAGTATTTGAAGACCGCT 3091
Qy 3016 CAGACACTGATGTGTTCAGAAATAGGTGAACGTCTGCGCGAAGCTTTAAGCCA 3075
Db 3092 GCGCTGTGACGCGGTTCAGAAACAAAGCAACGTGTGTGGCGAAAGCCTGTGCT 3151
Qy 3076 GTCTGGCCACGGCCAACTTGTCTGACGAGACAGACGAGTGGAGAC---GTGACCCA 3132
Db 3152 GTCTTGACACTGCGGATAGATTGACAGACAGAGGTGAGACCACTAATTAACAGCA 3211
Qy 3133 TTCAAGCATGACAGAGGTACTCACTGAAATGGCACTGAACCTTTTGGACCAAGTTC 3192
Db 3212 TTTAAGGAGGACAAAGTTACTCTCAGTGGTGGCTTGAATTTGACCAAGTAC 3271
Qy 3193 TTTGAGTAGACTGACAGTGGGTTATTTTCCGCTCTACCGTGCACCTTAATTACAG 3252
Db 3272 TATGAGTTGACCTGACAGTGGCGCTGTTTCTGCCGGAAGTGTCCCTGTATTAACG 3331
Qy 3253 GATAGACCTGGGATTAATCTGCCAGGGAAGAACTGTATGGGCTTAATAGAGAGTGA 3312
Db 3332 AACCAACACTGGGATTAACAGACTGTGTGAAGATGTATGAATTCAGCGCAACGT 3391
Qy 3313 AAGAGTTGTACCGGATATCCGTGCATCAAAAAGCGGTTGACACAGCAGGAGTACT 3372
Db 3392 GCCAGCTGGAAGCTAATACCTTCTGAAAGGCGAGTGGCATCCGCGAACAGGCA 3451
Qy 3373 GATATTAAGAAATTAACCATCAAGGACTACTCTCAACATTAATGTGTTCATTAAT 3432
Db 3452 GTTATCCAGAAAATAATCCAAACCGCTTCTGTGTGACAAATGTATTCATCAAC 3511
Qy 3433 CGCCGGTTGCCCACTGTTATCGTTGACCAAAAGACAGGTAACAATGATCACAGC 3492
Db 3512 CGCAGGCTGCGACGCGCTGTGTGAGTACAAAGCGTTAAAGCAGTAGAGGTGAG 3571
Qy 3493 GGATTTCTATTAAGATGAAGGCAAAATCTGTGTGTGATCGGCGATCTATCAGACT 3552
Db 3572 TGGGTGCTATTAAGTAAAGAGGTAACAGTCTGTGTGATGAGTAACTGTGCT 3631
Qy 3553 CCAGGGAAGAAAGTAGTCCATGGGTCATTTGCCACTAATAC-----CATCAG 3603
Db 3632 TTGCTCGACGACAGGGTCACTGTGTGTGACCGCTGAATGCACAGGCGCGATAGGTC 3691
Qy 3604 TGTGATCTCGATTTGGGAATACCTAGCAATGCGTAAATATGACATTAATCTTTGCAAT 3663
Db 3692 TACACCTTAAGTTTAGACTCCGCTGACCGCGCAGGTTGCACTGTGTGTAAC 3751
Qy 3664 GTTAGAGCCCGCTACAGAACCATCACTAACCAAGTGCAGAGTACGCTATCCACAC 3723
Db 3752 ATTACACGGAATTCAGAAATCCACATCAACGAGTGTGTGACACGCGCATGAAGTG 3811
Qy 3724 AGCATGCTAACGTGAAGGCTGTCCACCACTGAACATGTCGGAACAATGTGTGCTATA 3783
Db 3812 CAGATGCTTGGGGAGATGCGCTACGACTGTAAAAACCCGCGCGACTCT---TGATGAG 3868
Qy 3784 GGGATGCGGCTTGCTGATGCGGCAACCGAGATATATCATCTGGGGTGGCAGCCTCATTT 3843
Db 3869 GCTTACGGATACCGCTAATTAATACGGAAGCCGTTGTTCTTCTTAACGAAAGTTTC 3928
Qy 3844 AGGTTTACCCGTGTCTGTCAAGCTTAAGAACATGCGGAAAATACGTAGGTTCTTGTG 3903
Db 3929 TCGTGTCAAGAGATGTTGGCCCGGATGTGTGTCAACGAAATACAGAAAGTGTCTTGCTG 3988
Qy 3904 TTCTTGGCAGAGCAACCGCAACACACACATGACCGAGACAGATCGGTGTAGTCTT 3963
Db 3989 TTCTTCAACTTTTACCAACGAAAGAGACCTTCACTGCTACCAAGATGAATACCAACTG 4048
Qy 3964 GACAACTATATCAAGGGTCAACAGGTAAGAGGACAGAGAGAGCTCCAGGTAAGAGTG 4023
Db 4049 AGTCCGTGTATGCCGGAAGAACATGACACAGCGCGGGTGTGACCATCTACAGGTT 4108
Qy 4024 ATCAGAGTGAATTTAGCAAGCCGCTGACCAAGCTATCGTTAATGCTTAATAGCAAA 4083
Db 4109 AAGAGACAGACATAGCCCACTGCAACAGAGCGGCTGTGTTAAGCAGCTTAAGCGCCGT 4168
Qy 4084 GGTCAACAGGTTCCGGAAGTGTGCGGTGACCTGTACCGAAAATGCGCGGCTTTTGAT 4143
Db 4169 GGAATGTAGGGAGATGCGGTATGACAGGCGGTGCGAAGAAATGCGCGCTTTAAG 4228
Qy 4144 AGACAGCAATATGCTGTGCGGAGCGGTAGACTTGTGAAGCAGCAACCG---CTCATATA 4200
Db 4229 GGAAGCAACACAGTGGGCAATTAATAAGTCAATGTGCGGTGTATCCCGTATC 4288
Qy 4201 CATGCTGAGAACCCAAATTTTCTTAAGATCCGGAACCGGAGGGCGACTTAAGCTGCA 4260
Db 4289 CAGCTGTACGCTTAATTTCTGTGCAAGCTGAAGCGGAAGGGGACCCGTAATGGCC 4348
Qy 4261 GCTGCTCAATGAGACTAGGCTCATGTCAACCGCTGAGCGGATTAACAAAATATCATGA 4320
Db 4349 GCTGTCAACGGGAGAGTGGCGCGCGGAATGAACAGACTGTCACTGACACAGTACCATC 4408
Qy 4321 CCGCTATCTGCAACCGGCTATTTCTGTGTGGCAAGATGAGTATGATCATATGCTCAT 4380
Db 4409 CCGCTGTGTTCACAGAGATGTTTCAAGGCGGAGAGATAGCTGTGACGATCCCTCAAC 4468
Qy 4381 CACCTGTTCATGCTTTTGCACATGAGATGCGATGCTCAATATATTTGCTGGATAAA 4440
Db 4469 CATCTATTTACAGAAATGAGCGCAGAGCGTGAACGTATCTCACTCAAGACAA 4528
Qy 4441 CAATGGAGACACAGATATGAGAGCCATTCAACCGCAAAAGAACGCTGCAAAATTTCTGAT 4500
Db 4529 AGTTGGAGAGAAATAATCCAGGAAGCCATTGACATGAGAGCGGTGTGAGTGTCAAT 4588
Qy 4501 GATGACAAAGCAGTACATTTGCTGTAGGGTCCACCACAAACAGCTCTTTGGCAGGC 4560
Db 4589 GATGACGTGAGCTGACCAACAGCTTGTGTAGAGTGCACCCGAGACAGCTGTGTGGT 4648
Qy 4561 AGACAGATTACTCCGTCAATGAGAGCAAGTTGTATCATACCTGAAAGTACACGATTC 4620
Db 4649 CGTAAAGGCTACAGTACCACTGACGCGGTGCTGTACTGTAAAGGTAACGAAATTC 4708
Qy 4621 CATGACCGCCAGAGACTTGTCCGAAATTCATGACAAATGTGGGCCCAAAATCTGAGGCT 4680
Db 4709 AACCAAGCTCTATGATTAATGAGAGATGCTGACGTTGTGGCCCAACCTGCAAGAGCA 4768
Qy 4681 AATGAGAGATTTGCTGTATCATCTGTGGGGAAGTATGTCTACAGATCCGCTCCAAATGC 4740
Db 4769 AACGAAACGATATGCTTATGCTGCGCTGTGGGAAACAAATGACAAATCAATCCAAATGT 4828
Qy 4741 CCAATAGAGAGTCAAGAGCGTGTGCTCACTTCAACACTTCCATGCTGTGTATTAAC 4800
Db 4829 CCGGTGAACGATTTCCGATTCATGAACCTCCAGAGACAGTGCCTGCTGTGCGCTAC 4888
Qy 4801 GCTATGACGGTGAAGCGCGTATACAGTGTGCGCTGTGCAAGAAAGAAAGTGTGCGCTGA 4860
Db 4889 GCAATGACAGACAGAACGATGTGCGCGCTTAAGTCAACACAAATTAAGCAATGTGTGT 4948
```

Qy 4861 TGTTCATCATCTCTGTTGGCCGAGTACAGATCACAGCGCTGCAGAACTACAGTGCAGC 4920
Db 4949 TGTTCATCATCTCTGTTGGCCGAGTACAGATCACAGCGCTGCAGAACTACAGTGCAGC 5008
Qy 4921 AAACCAATGCTGTTTTCAGAGCGCTGTTACACCGCTGTACACCCGAGAAAGTACCGGAA 4980
Db 5009 AAGGTTCTCTGTTGACCGCGAGGTAAGTACCTTCAAGTGTAGTCCCGGAAAGTATCCGCA 5068

RESULT 11
AAK76582 standard; DNA; 15538 BP.
AAK76582;
AAK76582;
11-AUG-1999 (first entry)
Plasmid pMP44 nucleotide sequence.
Plasmid pMP44; hepatitis delta ribozyme; alphavirus vector; vaccine;
paramyxovirus; respiratory syncytial virus fusion protein; RSV F;
detection; infection; immune response; immunisation; ss.
Synthetic.
MO9925858-A1.
27-MAY-1999.
13-NOV-1998; 98WO-CA01064.
14-NOV-1997; 97US-0065791.
(CONN-) CONNAUGHT LAB LTD.
Li X, Parrington M;
WPI, 1999-370764/31.
Improved alphavirus vector containing DNA encoding a paramyxovirus protein
Claim 19; Fig 3; 61pp; English.
The present invention describes an improved alphavirus vector containing DNA encoding an immunogenic paramyxovirus protein and DNA to enhance the immunoprotective ability of the paramyxovirus protein in vivo. The vector comprises: (1) a first DNA sequence which is complementary to at least part of an alphavirus RNA genome and including the complement of the complete alphavirus RNA genome replication regions to permit in vivo replication; and (2) a second DNA sequence encoding a paramyxovirus protein or protein fragment that generates antibodies that specifically react with the paramyxovirus protein; where the second DNA sequence is inserted into a region of the first DNA sequence which is non-essential for replication, the first and second DNA sequences being under transcriptional control of a promoter. The vector is used to immunize a host against disease caused by infection with paramyxovirus, especially respiratory syncytial virus. Human RSV is a major pathogen responsible for severe respiratory tract infections in infants, young children and the institutionalized elderly. The vector provides a protective immune response using a lower dose and less time than vectors in US8476397 and US8896500. Additionally the vectors using native RSV F produce protective immune responses in the absence of pretreatment of the animal model with cardioloxin, a material known to increase the uptake of DNA and enhance the immune response. The present sequence represents a specifically claimed plasmid pMP44 nucleotide sequence from the present invention.
Sequence 15538 BP; 4313 A; 3774 C; 3837 G; 3614 T; 0 other;
Query Match 31.3%; Score 1564; DB 20; Length 15538;
Best Local Similarity 58.3%; Pred. No. 0;

Matches 2904; Conservative 0; Mismatches 2040; Indels 36; Gaps 8;
Qy .31 AGAATTCAGTGGCTTATGATGCTACAGCCCGTATGCAAGTGGTTCAGCGGCGTTT 90
Db 2642 AAAGTGCATGTTGATATGAGCTGACAGCCCATTCATCAAGTCTTTCAGAAAGGCAATT 2701
Qy 91 CCACATTTTGAATCGAAGCAAGGAGTCACTGCAATGACATGCAATGCCAGAGCG 150
Db 2702 CCGTGTTCGAGTGGATCATTTGCAAGTCAACCAATGACATGCAATGCCAGAGCA 2761
Qy 151 TTTTCGATGTTGGCAACCAAGCTCATTTAGAGCGAAGTGCACCGGACCAAGTATCTTG 210
Db 2762 TTTTCGACCTGGCTTACCAATGATGATGAGGAGAGCTGACCAAGACACACTCTCTTG 2821
Qy 211 GACATTTGAAGTGGCCCGTTCAGACATGCAATTCATTCACCTGATCTGATCTGCG 270
Db 2822 GATATCGGCAATGCGCTTTCAGAGAAATGATGTTGACCAATATCACCTGCTGATGCG 2881
Qy 271 CCTATGATTAAGCGCTGAAAGACCCGAGACAGCTACACCGGTATGCAAAAGACT-----T 324
Db 2882 CCTATGCGGACGCGAGAAAGACCCGAAAGGCTCGATGACTACGAAAGAACTGGCAGCG 2941
Qy 325 AAGAAAGTGAATTAACCGAAGAACTAGCTCTTAAGCGGCGAGACTGCTGGAAATC 384
Db 2942 GCTTCGGAAGAGTGGCTGATGATGAGAGATCGCAGAAATATCAACGACCTGCAACCGTC 3001
Qy 385 ATGTCAACACCAAGACGAGAACTTCATCTGTGTATGACACAGAGCCGCTGTAGG 444
Db 3002 ATGGCTACGCAAGCGCTGAATCTTCATCTTGTGCTGATACAGAGATGATCGTGTGT 3061
Qy 445 TACTTTGGAAGTGAAGAGTATTAACCAAGATGATGAGGATGATGACCGAGCAATTC 504
Db 3062 ACGGACCGGAGAGTGGCGCTGATACAGAGACCTGATGATGATGATGATGATGATGATG 3121
Qy 505 TACCAACAGCGGCTTAAAGAGTGAACAATTAAGTGAAGCTTTGACACGACCCCT 564
Db 3122 TACCATGAGCGAGTAAAGTGTGCAAGCGCGATGATGATGATGATGATGATGATGATG 3181
Qy 565 TTTATGTCAAAACATGAGGAGTTCCTACCTTCTTCAACACGAACTGGGCTGACGAG 624
Db 3182 TTTATGTTGACGCGCTGACAGCGCGCTATCAACCTGACCAAACTGGGCGACAG 3241
Qy 625 AGAGTATGGAAGACGATTAATGAGCTGCTCGTAACTCAGATCTTCAGAGAGCGGCTT 684
Db 3242 CAGGTGTACAGCGCAGGAACATAGACTGTGTGACAGATCTTCAGTGAAGAGATC 3301
Qy 685 GAAAACTCTCAATCTTGAAGAGAGGCTCAACCTACTAATTAAGATATATTCG 744
Db 3302 GCGAACTGTCAATTCCTCGCAAGAGCAATTTGAACCTTGCGACACAGTATGTTCTG 3361
Qy 745 GTTGTTCAACATCTACACAGAAATGATCACTGTTAGTACTGCTGCACTTCCAAAC 804
Db 3362 GTAGATCTTACATTTGACACAGAGCAAGAAAGTACTGAGAGCTGCACTTCCCTCC 3421
Qy 805 GTGTTCCACTGAAAGAAAGTAACTTAAGTCAAGTATGATGAGGAGCAATTTGTCAGTGT 864
Db 3422 GTATTCACCTGAAAGTAAAGTAACTTTTACCTGTAGTGTGATGATGATGATGATGATG 3481
Qy 865 GAAGGATGATCATCAAAAGATTAAGATCAAGTCAAGGACATTAAGTGAAGTGAAG 924
Db 3482 GAAGGATGATCAAGTAAAGATTAAGATCAAGTCAAGGACATTAAGTGAAGTGAAG 3541
Qy 925 TTGGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 984
Db 3542 TACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3601
Qy 985 GCGGAGGAGGTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1044
Db 3602 GGAAGAAAGATCTATTCCTGTATGCACTTACCTGCTTACCAATGATGATCAATG 3661
Qy 1045 ACAGGATTTGAGCAATGACGTTAGTGTGATGATGATGATGATGATGATGATGATG 1104
Db 3662 ACTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3721

1105 AACCAAGATGTCGTCATGTAGAGCGCAAGAAATACTTAACAAATGACAGAT 1164
1165 CTATTACAGAGTGCCTCCAGGCGTTTTCAGGTGGCGCGTGAACATCGTCCGACTTG 1224
3782 CTGCTTCGATGTGGCGGTGCGATTTAGCAAGTGGGGAAGAAATCAAGGACGACTT 3841
1225 GAGCAGCAAGAAAGAACTAGGGGTGGGAGCGCACTTTACTATAGGGCTGCTGGGCT 1284
3842 GATATATAAAACTCTGGGGGTGCGAGAGAGTCACTTACTTCTGCTGTGGGCA 3901
1285 TTCAAGCCCAAGAAATACATCCATCTACAAAGAGCTGTAGCGAAACAATTAGAA 1344
3902 TTTAAAAAGAGAAAGATGCAACATGTAAGAAACAGACACCCAGCAATGTGAG 3961
1345 GTACCTGCGCTTGTGATCTATTTGTGATTCACGCGCTTACAGCCAGCGGCTGATATG 1404
3962 GTGCTTCAAGGTTTAACTGCTGTGATCCGAGCGTATGTTACAGGCGCTGCAATC 4021
1405 GCGTTCGCGCTAGCTCAGCTGCTGTTGAACCACTGTCAACCCGACCGGCTATT 1464
4022 CCAGTCAGATCACCATTTAAGATGCTTTGGCCAAAGAACCAAGC--GAGAGTTATA 4078
1465 ACAATGGCCGATGGAGCATCTGGGCGTTACAGCAAGAAAGCTGAAGAGTGGTGA 1524
4079 CTGTTCTCGAGCGGTGCTACGCCAGGAGTGTGTAACAAGAGAGAGAGTGGAG 4138
1525 GCGGAGAGATCAGAGAGCCCTGCAACCTTGTCTCCCTGAATAAGAAAGAACGCTA 1584
4139 GCGGAGTGAATAAGAGCTTACACCTCTGCTCCGATGCGCGCGGAGACGGGA 4198
1585 GAGCAGAGATGACCTCATTAATGCAAG--GAGAGAGAGTGAAGTGAAGACA 1638
4199 GTCTGCACTGAGCTTGAAGAACTAGAGATGACAGGAGTGAAGGCTGTGAACA 4258
1639 CCAGAGAGACATCAGGGGTGACAAAGTTACCCAGCGAAGAGATTGGGTCTTACGCT 1698
4259 CTTGCGAGCGGTGAAGTACCGGACAGCGCAACGATCTACTAGAAATTTAGCTA 4318
1699 ATACTTTCAACCCAGCGGTATTGAATAGTAAAACTGGGCTGATCCACCAATGGCG 1758
4319 GTTCTGCTCCCGAGACCGTCTCAAGAGCTCAAGTTGGCCCCCGTACCTCTAGCA 4378
1759 GAACAGATCTGTAATGATCTCAAAAGTGGGAGGAGATCAAAAGTGAAGCTATC 1818
4379 GAGCAGGTAAAAATATTAACATTAACGAGGCGCGGCTTACAGGTGACGAGATAT 4438
1819 CACGTAAGGTCAATGTATCAGAAAGGAGCGGCGCTCTGTTCAAGACTTCCAGGCAATTG 1878
4439 GACGCGAGGCTCTACTACATGTGATCGGCAATTCGAGTCCCTGAGTTTCAAGCTTTG 4498
1879 AGTGAAGCGCTACGATCGTTTTCACAGAGAGAGTGTCTTAAACAGATCTCCAGCAC 1938
4499 AGCAGAGCGCCTATGTGTAACAGAAAGGAGTTCCTCAAGAGAACTATACAT 4558
1939 ATCCGATCAACGAGAGCGCTTAACTGACAGAGATCTATTAAGCTGTAAAGACT 1998
4559 ATTCCGTTCAACGAGCGCTCTGTAACACGAGAGAGAACTACGAAAGGCTAGAGCT 4618
1999 CAGACACAGCTAGATAGTCTGATTTGACGAGAGAAAGTGTAAACGAGAA 2058
4619 GAAAGAACTGAGCGGAGTACGTGTTGACGTAAGTAAAAATGCTGTCAAGAGAGAG 4678
2059 GACGAGGTCCTTGTGCTTAACCGGTGATCTGTAGATCCACATTTCAAGAGTTGCG 2118
4679 GAAAGCTGCGGTTTGTGTGTGGAGAGCTTAACCAACCCCGTTCAGATTTGCGC 4738
2119 TAGCAGATCTCAAGACAGACAGACAGCTTCAAAAGTCCCAACCATGAGATCTAT 2178
4739 TACGAAGGCTGAAGATCAGGCGCTGCGACCATATTAAGACTAGTAGAGTCTTT 4798

2179 GAGTCCAGAGTTCAAGTAATCTGAAATCATCAAAAGCGCTGTGACTAAGAAAGATCTG 2238
4799 GGGGTTCCGGGATCAGGCAAGCTGTGATTTAAGAGCGCTCGTACCAAAACGATCTG 4858
2239 GTTGTAGTCCGAAGAAAGAACTCGGCAAAATCATAGGATGTAGAGAGATGAGA 2298
4859 GTCAACGCGGCAAGAGAGAACTGCAAGAAATAGTTAACAACGTAAGAAAGACCGC 4918
2299 CGTATGATGTTCTGCTAGAGCTGTGATTCAGTGCTGCTTAATAGGGGTTAAGACCCC 2358
4919 GGGAGGGGCAAGTAGGAAACAGTACTTCATCTGCTTAAAGGGGTCTGTGCGC 4978
2359 GTTAACATCTGTACATTTGATGAGGCAATTCCTGCAATCAGAGAGCGTGTGCACTG 2418
4979 GTGACATCTCTATATGTGAGAGAGGCTTTGCTTGCATTCGGTACTGTGCGGCTTA 5038
2419 ATTGCACTGCTCAACCT--AAGAAATGATATTGTGCGGGAACCCAAACAATGCGC 2475
5039 ATTGCTCTGTTAACTCGAGCAAAAGTGTGTATGCGGAGACCCCAAGCAATGCGGA 5098
2476 TTTCTTAACTGATGTGCTGAAAGTACATTTTAACATGACATPAGCAATGATGATAC 2535
5099 TTTCTCAATATGATGACGCTTAAAGTAACTTCAACCAACATCTGCACTGAATGTGT 5158
2536 CATAAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATGCTCTCACGCTCTTC 2595
5159 CATAAAGTATATTCAGACCTTACGCGCTGCACTGACGCGCATCGTGTCTAGGTTGAC 5218
2596 TACGACAGCGAATGAAGAGGTTAAACCATGTGCTGATTAATCATATGATCCACA 2655
5219 TACGAGGCAAGATGCGCACAGACCAACCCGTGCAACCAACCAATATCATAGACACACA 5278
2656 GGAACCAAGAGCGGCAAAAGATGATCTGATTTCTAATCTGTTTACAGAGGTTGAAA 2715
5279 GAGACAGCAAGGCCCAAGAGACATGCTTAACTGATCTTCCAGGCTGGGCAAG 5338
2716 CAGCTACAGATTGACTCAAAATAATCAAGAAATGATGATGCTGCGCTGATCGAAGACTT 2775
5339 CAGCTGAGTTGGATCTACGCTGACAGAGTATGATGACAGAGAGATCTCAGGGGCTC 5398
2776 ACGGGAAGGCGTTTATGCTGTGAGTACAAAGTCAACAGAAATCCACTCTACTGCGAG 2835
5399 ACCGCAAGGGGATACGCGGTAAAGCAGAGTGAATAAATCCCTGTATGCCCCCT 5458
2836 ACTTCTGAGACGGAACGTTACTTACTTACAGGACAGAAAGCAATGTCTGGAAGAG 2895
5459 GCGTGGAGCGTAATGTATCTGTACCGGCACTGAGATAGCTGTGTGGAAGAACG 5518
2896 CTAGCTGTATCCCTGATTAAGACATTAACGCTAATATATCCCGGAGATTCAAGGCT 2955
5519 CTGGCGGCGATCCCTGATTTAAGGTCTTATCAAACTTCCACAGGTTACTTACGGCC 5578
2956 TCATTGAGCACTGACGCGGAAACAGACGCCATTAATGCAAGCGCTTCTTGAAGCCG 3015
5579 ACATTGGAATAATGCAAGAGAAACAGCAAAATATGAAGTGAATGAAGACCGGCT 5638
3016 CAGACAGCTATGTGTTCAAGATTAAGGTGAAGTCTGCTGGGGAAGCTTTAGAGCCA 3075
5639 GCGCTGTGAGCGGTTTCAAGAACAAAGCAAGCTGTGTGGGCAAAACCTGTGTGCT 5698
3076 GTCTTGGCAACGGCAACATTTGTGCTGACAGACAGAGTGGAGAC--GTTGACCCA 3132
5699 GTCTGGAACACTGCGGAATCAGATTGACAGAGAGAGTGGAGACCATATTTACAGCA 5758
3133 TTCAAGCATGACAGAGGTTACTCACTGAATGGCACTGAATCTTTTGTGACCAAGTTTC 3192
5759 TTTAAGAGAGACAGAGCTTACTCTCAGTGTGCTTGAATGAATTTTCAACCAAGTAC 5818
3193 TTTGGAATTAACCTGGAACAGTGGTTATTTTCCGCTCTTCCGTCGACTTAAGTACAGG 3252
5819 TATGAGATTAACCTGGAACAGTGGCTGTGTTCGCCCCGAAGGTGCTCTGTATTAAGAG 5878
3253 GATCAGCACTGGATTAATCTGCCAGGGAAGAAATGATAGGCTTAATAGAGGTAGCA 3312

```

Db      5879  AACAACCACTGGGATTAACAGACTGTGGAGAGATGATGATCAATCCGCAACAGCT 5938
Qy      3313  AAGAGTGTGTCACGGCGATATCCGTGCATCAAAAAGCGTTGACACAGGCGGTAGT 3372
Db      5939  GCCAGCTGGAAAGTAGACATACCTTCTGAAGGGCGAGTGGCATACGGGAAGAGCA 5998
Qy      3373  GATATAGGAATATATACATCAAGAGCTACTCTCCAAACAATTAATGTGGTTCCATTAA 3432
Db      5999  GTTATCGGAGAAAGAAAATCCACCGCTTCTGTCTGGACAAATGTAATTCCTATCAAC 6058
Qy      3433  CGCCGTTGCCCTACTCTGTGATCGTTGACCAAAAGACAGGGTACAACTGATCAACG 3492
Db      6059  CGCAGGCTGCCACAGCCCTGTGGCTGAGTACAAAGACGGTTAAAGCGAGTGGGTGAG 6118
Qy      3493  GGAATCTCATCTAATGATGAAGGCAATCTGTGTGTGATCGGGATCTCTATCAGCAT 3552
Db      6119  TGGCTGTCAATTAAGTAAAGAGGGTACCACTGTCTGTGTGATGATGATCAAACTGGCT 6178
Qy      3553  CCAGGGAAGAAAGTAGAGTCCATGGGTCCATTGCCCATTAATAC-----CATCAGG 3603
Db      6179  TTGCCTCGACGAGGATCACTTGGTTGTCTACCGCTGAATGTCAAGGCGCGATAGTGC 6238
Qy      3604  TGTGATCTCGATTTGGGAATACCTAGCAATGTGGTAAATATGATTAATCTTTGTCAAT 3663
Db      6239  TACGACTTAAGTTTAGACTGCGGGCTGACGCGGAGGTTGCACTTGGTCTTTGTGAAC 6298
Qy      3664  GTTAGAGACCCGCTACAGAAACATACCTACCAACAGTGGAGATCAAGCTATCCACAC 3723
Db      6299  ATTCAACAGGAATTCAGATTCACCACTACAGAGTGTGTGACACGCGCCATGAAAGCTG 6358
Qy      3724  AGCATGCTAACGTTAAGGCTGTGCCACCACTGAACACTGGGGGAACAATGTGTGCTATA 3783
Db      6359  CAGATGCTTTGGGGGAGATGCTGCTAAGCTTAACAACCGGCGGCACTCT--TGATGAGA 6415
Qy      3784  GGGTATGGGCTTGTGATCGCGCAACCGAATATATCATCTGCGGTGCAACGCTCATTT 3843
Db      6416  GCTTACGATACGCGATTAATTAACGCAAGCGGTTGTTCTCTTAAGCAAGAAAGTTCC 6475
Qy      3844  AGTTTACCCGTGTCTGTACAGCTTAAGAACATCGCCGAAATATCTGAGTTCTTTCGTG 3903
Db      6476  TCGTGTGAAGAGTGTGGCCCGGATGTGTGACACAGCAATACGAAGTGTTCCTTGGCTG 6535
Qy      3904  TTCTTCGGCAAGGACCAAGGCAACACACATGACAGAGAGACTCGGTGTAGTGGCTT 3963
Db      6536  TTCTCAACTTTGACAAAGGAAAGAGACCTCTTACGCTTACACAGATGATACCAAGCTG 6595
Qy      3964  GACCAATCTATCAAGGTCACACAGTACAGAGGAGGAGAGCTTCAAGCTTACAGAGTG 4023
Db      6596  AGTCCGTGTATGCGGAGAAAGCATGACACGCGCGGTGTGACATCTTACAGAGTT 6655
Qy      4024  ATCAGAGGTGACATTAAGCAAGCGCTGACCAAGCTATCTGTAATGTCTTAATAGCAAA 4083
Db      6656  AAGAGAGGACATATGCGACAGTGCACAGAGCGGCTGTGTAAAGCAGCTTAAACCCCGT 6715
Qy      4084  GGTCAACAGGTTCCGAGGTGTGGGTGCACTGTACCAAAATGCGCGCTGTTTAT 4143
Db      6716  GGAATCTGAGGGAGTGGCTATGACAGGCGGTGCGAAGAAATGCGCGCTTAAAG 6775
Qy      4144  AGACAGCCAATAGCTGTGCGGACGCGCTAGACTTGTGAAGACGAACCG---CTCATCTATA 4200
Db      6776  GGAGACGACAAACAGTGGGCAATTAACAGTCAATGTGGCTGTGATCCCGCTCATC 6835
Qy      4201  CATGCTGTAGAACCAATTTTCTTAAGATCGCGGAACCGGAGGCGACCTTAAGCTCGCA 4260
Db      6836  CACGCTGTAGCGCTTAATTTCTCTCCAGTGAAGCGGAAGGGGACCGCAATTTGGCC 6895
Qy      4261  GCTGCTCATGAGATAGCGTTCATCGTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db      6896  GCTGTCTACCGGGCAGTGGCGCGCAAGTAACACATGTCTACGTAGCGCTTACCATC 6955
Qy      4321  CCGCTACTGTCAACCGGATCTTATCTGTGGCAAGATCGAGTATCATCTGATCAT 4380

```

```

Db      6956  CCGCTGCTGTCCACAGAGTGTTCAGCGCGGAAGAGATAGGCTGACAAATCCCTCAAC 7015
Qy      4381  CACCTGTTCATGCTTTTGACATCTAGCATGCGGATGCGGATGACCATATATGCTTGATAA 4440
Db      7016  CATCTATTACAGCAATGAGAGCGCCAGGACGCTGACGTACCTACTACTGACAGACAAA 7075
Qy      4441  CAATGGAGACCAAGATTAATGAGGCGCATTCACCCCAAGAAAGCGTCAAAATTTCTGAT 4500
Db      7076  AGTTGGAGAAAGAAATTCAGAGAACCATTTGACATGAGAGAGCGCTGTGAGTTGCTCAAT 7135
Qy      4501  GATGACAAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4560
Db      7136  GATGACGTGAGAGTACCAAGATGATGATGATGATGATGATGATGATGATGATGATG 7195
Qy      4561  AGACAGGTTACTCCGTAATGAGGCGCAAGTTGATGATGATGATGATGATGATGATGATG 4620
Db      7196  GATAAGGCTACAGTACCACTGACGCGGTGCTGTGATGATGATGATGATGATGATGATG 7255
Qy      4621  CATCAGACCGCCAAAGACATTTGCCAAATCCATGCAATGTGGCCCAACAAATCTGAGCT 4680
Db      7256  AACAGGCTGCTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7315
Qy      4681  AATGACAGATTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4740
Db      7316  AACGAAAGATATGCTATATGCGCGTGGCGAAACAAATGACAAATCAGATCCAAATGT 7375
Qy      4741  CCAATGAGAGATGAGAGCGCTTGTCTTCACTTCAACATCTTCACTGCTGTGATTAATC 4800
Db      7376  CCGGGAACGATTCGATTCATCAACACTCCAGAGAGTGGCTGTGCTGTGCTGTGCTGTG 7435
Qy      4801  GCTATGACGCGTGAAGCGGTATACAGTGTGGCTGTGCAAGAAAGAACATGTCGCGTA 4860
Db      7436  GCAATGACAGAGAGAGATGCGCGCTTGTGATGATGATGATGATGATGATGATGATGATG 7495
Qy      4861  TGTTCATCATCTTCTGTTGCCGAATGATGATGATGATGATGATGATGATGATGATGATG 4920
Db      7496  TGTTCATCTTTTCCCTCCGAAATATCAGATGATGATGATGATGATGATGATGATGATG 7555
Qy      4921  AAACAGTCTGTTTTCAGGCGTGTATCACCGGCTGTATCACCCGAGAAATGATGCGGAA 4980
Db      7556  AAGTTTCTCTGTTCGACCCGACGATCTTCAAGTGTGATGATGATGATGATGATGATG 7615

RESULT 12
ABN86687
ID ABN86687 standard; DNA; 11489 BP.
XX
AC ABN86687;
XX
DT 05-NOV-2002 (first entry)
XX
DE Nucleotide sequence of a pSCA1 suicide DNA vector.
XX
KW Major histocompatibility complex; MHC; antigen presenting cell; APC;
XX antigen; cytotoxic; virulence; gene therapy; CD8; vaccine; therapeutic;
XX cancer; viral infection; ds.
XX
OS Synthetic.
XX
PN WO200261113-A2.
XX
PD 08-AUG-2002.
XX
PF 01-FEB-2002; 2002WO-US02598.
XX
PR 01-FEB-2001; 2001US-26534P.
XX
PA (UNIO ) UNIV JOHNS HOPKINS.
XX
PI Wu T, Hung C;
XX
DR WPI; 2002-619261/66.
XX

```


1832 GAGCGAGTGAATAATATACATACATACGCGAGGCGGCGGTATACAGGTCCAGCGAATAT 1891
1819 CACGGTAAAGTCATTTGATGACACGAGGAGCGGGGTCCTGTTCAAGACTCCAGGCAATTG 1878
1882 GACGGCAGGGTCTCTACTACATGATGATGCGCATTCGGTCCGTGAGTTCAACCTTGG 1951
1879 AGTGAAGCGCTACGATCGTTTTCACGAGAGGAGTGTGTAAGAAGATACCTGAGCAC 1938
1952 AGCGAGAGCGCCACTATGCTGTACACGAGAGGAGTTCGTCAACGAGAACTATATCCAT 2011
1939 ATCGCATATACGAGAGAGCGCTAAACACTGACGAGAGTACTATPAAGCTGTAAAGACT 1998
2012 ATTGCGTTCAACGAGCGCGTCTGTAACACCGAGAGAGAACTACGAGAAAGTCAAGAGCT 2071
1999 CAGGACAGAGCTCAGAAATACGTCCTGATATGAGCGACGAAAGTGTGTAAGGAGAA 2058
2072 GAAAGAACTGAGCGCGAGTACGTTGTCGAGTAAATAAATGCTGCGTCAAGAGAGAG 2131
2059 GACGAGGTCCCTTGTGCTTAACCGGTGATCTGTAGATCCACATTTTCACGAGTTGGC 2118
2132 GAAAGGTGGGGTTGTGTGTGGTGGAGAGCTAACCAACCCCGCTTCATTAATTCGCC 2191
2119 TACGAGACTCTCAAGACAGACGACGACGACCTCAAAATGCCAAACCATCGAGCTTAT 2178
2192 TACGAAAGGCTGAAATGATGAGCGCGTCCGACCATATPAAGACTACAGTAGTAGAGTCTT 2251
2119 GAGAGCGCAGGTTCCAGTAAATCTGAAATCATCAAAAGCGCTGTACATPAAGAAATCTG 2238
2252 GGGGTTCCGGGATCAGGCAAGTCTCTATTTAAAGAGCTGTGTAACCAACCATCTG 2311
2239 GTTGTAGTGCAGAAAGGAAATCTGCGAGAAATCATCAGGAGTGTAAAGAGATGAGA 2298
2312 GTCAACACGCGAABAGAGAACTGCCAGAAATATGTTAAGAGCTGAAGAGAACGCC 2371
2299 CGTATGATGTTGCTGTAGAGCTGTGATTCAGTGTCTTAATGGGTTAAAGCAACC 2358
2372 GGGAAAGGGAACAAATGAGGAAACAGTGACTCATCTGCTAAACGGGGTGTGCTGTGCC 2431
2359 GTTAACTCTGTATATGATGAGGCAATTTGCTGCTCATGACGAGGAGCGTGTGCACATG 2418
2432 GTGACATCTTAATATGATGAGCGAGGCTTTCGTACCATTCGGTATCTGTGTCGCCCTA 2491
2419 ATTGCATCTGTCAAACT---AAGAAAGGTATTTGTGGGGAGCCCAAAACAATGCGGC 2475
2492 ATTGCTCTTTTAACTCTGAGGAAAGTGTGTTATGCGAGACCCCAAGCAATGCGGA 2551
2476 TTCTTTAACTATGATGCTGTAAGATGATATTTTAACTATGACATATGACATGAGTATAC 2535
2552 TTCTTCAATATGATGACAGCTTAAGGTGAATTCACACCAACATCTGCACCTGAATATGT 2611
2536 CATTAAGACATCTTAGAGAGTGCACACAGACTGTAAACCGCATGTCTCCACGCTTTC 2595
2612 CATTAAGATATATCCAGAGTTCGACGCGTCCAGTACAGCGCATGTGTCTGCTGTGAC 2671
2596 TACGACAGCGAATGAAGCGTTAACCCATGCTGTAATAAATGATATATATACCA 2655
2672 TACGAGGCAAGATGCGACGACCAACCGTGAACAAACCATATATATACACCA 2731
2656 GGGACCAAAAGCCGCAAAAGATGATGATTTCTTAACTGTTTCAAGAGATGAGTGA 2715
2732 GACAGACCAAGCCCAAGCGAGACATCGTGTAAATGCTTCGAGGCTGGGCAAG 2791
2716 CAGCTACGATTTGATCAAAAATATCGAATATGATGCTGCGGTGATCGCAAGAGATT 2775
2792 CAGCTGAGTTGATGATCGGTGACACGAAATGATGACAGCAGCAGCATCTCAGGCGCTC 2851
2776 ACGGGAAGAGCGGTTATGCTGTGAGTAAAGTCAACGAGAAATCCATCTACTGCGAG 2835
2852 ACCGCAAAAGGGATATAGCGCTTAAGGCGAAGGTGAATGAATAATCCCTTGTATGCCCT 2911
2836 ACTTGTGACAGTGAACGTGTATCTTACGACGACGAGAAACGCAATTTCTGAGAGCG 2895

2912 GCGTCGAGCAGTGAATGTACTGCTGACGCGCACTGAGATAGCGTGTGAAAAAG 2971
2896 CTAGCTGTGATTCCTGGAATAAGCACTTACAGCTAAATATCCGGGGATTTTACCGCT 2955
2972 CTGGCCGGCGATCCCTGATTAAGGTCTTATCAAACTTCCACAGGATCTTAAGGCG 3031
2956 TCATTGACGACTGCGACGCGGAACACGACGCCATTAATGAGCAGCGCTTGTGATPAAGCG 3015
3032 ACATTGGAAGATGCGAAGAAAGAACAGCAAAATTAAGAAAGTATGAAGAGACCGCT 3091
3016 CAGACGCTGATGTGTTCAGAAATPAAGGTAACTGTCTGCGCGAAGAGCTTTAGAGCA 3075
3092 GCGCTGTGAGCAGCGTTCCAAACAAAGCAGTGTGTTGGCGAAAAAGCTGTGCT 3151
3076 GTCTTGCGCAGCGCAATGTGTGTGACGAGACAGCTGGAGAC---GTGACACCA 3132
3152 GTCTGACACTGCGGAAATGATGACACGAGGAGTGAAGCATAATTAACCA 3211
3133 TTCAAGCATGACAGAGCTACTCACTGAAATGGAAGCACTTCTTTGACACGAGTTC 3192
3212 TTTAAGAGAGACAGAGCTTACTCTCAAGTGGTGGCTTGAATGAATTTGACCAAGTAC 3271
3193 TTTGAGTGAACCTGAGCAGTGGGTATTTTCCGCTCTTACCGTGCATTAATTACAG 3252
3272 TATGAGTGTGACCTGACAGTGGCTGTTTCTGCGCGAAGGTGTCCCTGATTAACGAG 3331
3253 GATCAGCATGAGATPACTGCGCAGGAGAAACATGTAAGGCTTAATPAAGAGGTACCA 3312
3332 AACACCACTGGAGTAACAGACTGTGTGAAGAGATGATTAATCAAGCGCAACACT 3391
3313 AAGAGTTGTACCGCGATATCCGTCATCAAAAGCGGTTGACACAGCAGGAGTACT 3372
3392 GCGAGCTGGAAGTACATGACATACCTTCTGAAAGGGCGAGTGGCAACGCGGACAGGCA 3451
3373 GATATPAAGAAATATATCATTAAGAGTACTCTCCAAATTAATGTTGCTTCAATTAAT 3432
3452 GTTATGCGAAGAAATAAATCAACCGCTTCTGTGCTGACAAATGTAATTCATATCAAC 3511
3433 CGCGGTTGCCCACTGCTTGAATGTTGACCAAAAGACAGGATCAACATGATACAGC 3492
3512 GCGAGGCTGCGCACGCGCTGTGTGATGATCAAGAGCGTTAAAGGACGATGAGGTTAG 3571
3493 GAATTCCTATCTAAGATGAAGGCAAAATCTGTGTGATGATCGGAGTCTATCAGACT 3552
3572 TGCTGTCAATTAAGTAAAGGATACCAAGTCTGTGTGTGATGATACCACTGGCT 3631
3553 CAGGGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3603
3632 TTGCTGCGACGAGGATCACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3691
3604 TGTGATCTGATTTTGGGATATCTAGCCATGTGCTGAATATGACATTAATCTTTGTCAAT 3663
3692 TACACCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3751
3664 GTTGAAGCCCGTACAGGAACATCACTATACCAAGTGTGAGGATCAGGATATCCACAC 3723
3752 ATTACACGGAATTCAGATCACCACTACCAAGTGTGATGATGATGATGATGATGATGATGAT 3811
3724 AGCATGCTTAAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3783
3812 CAGATGCTTGGGGAGATGCGCTAGACTGTAAACCCGCGGATCT---TGATGAGA 3868
3784 GGGTATGGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3843
3869 GCTTACGATGATGCGGATTAATCAAGCGAACCGTGTGTTCTCTTAAAGCAAGATTC 3928
3844 AGGTTTACCGGTGTGCTGACGCTTAAAGAACTGTGCGCAAAATGATGATGATGATGATGATGAT 3903
3929 TGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3988
3904 TTCTTGGCAGAGGACCAAGCGCAACCAACATGATGATGATGATGATGATGATGATGATGATGAT 3963
3989 TTCTTCAACTTTGACCAAGGAAAGAGACCTGTAGGCTACACGATGATGATGATGATGATGATGAT 4048

```

QY 3964 GACAACTTATCAAGGGTCAACAGGATGAGGAGGAGGCTCCAGGTACAGATG 4023
DB 4049 AGTCCCTGTATGCCGAGAAAGCCATGACACGGCCGGTGTGACCACTTACAGAGTT 4108
QY 4024 ATCAGAGGTGACATTAGCAAGAGCGCTGACCCAGCTATCTGTTAAATGCTTATAGCAAA 4083
DB 4109 AAGAGAGAGACATAGGACAGTGCACAGAAAGCGGTGTGTAAAGCAGCTAACGCCCGT 4168
QY 4084 GGTACACAGGTTCCGAGTGTGGGTGACTGTACCGAAATAGCGGCTGCTTTGAT 4143
DB 4169 GGAATGTAGGGGATGCGGTATGCAAGGGCCGTGCGAAGAAATAGCCCTTAAAG 4228
QY 4144 AGACAGCAATAGTGTGGGACGGCTAGCTTGTGAAGACGAAACG---CTCATATA 4200
DB 4229 GGAAGACAGACACCAAGTGGGACCAATTAACAGTATGTGCGGCTGTATCCCGTATC 4288
QY 4201 CATGCTGAGGACCAATTTTCTTAAGATGCCGGAACCGAGGGCGAACCCTTAAGCTGCA 4260
DB 4289 CAGCTGTAGCGCTAATTTCTGTCCACGACTGAAGCGAAGGGGACCGGAATTGGCC 4348
QY 4261 GCTGCTTACATAGACATAGCTCCATGCTCAACGCTGAGCGGATTAACAAATATCACTA 4320
DB 4349 GCTGTCTACCGGGAGTGGCCGCGAAGTAAACAGACTGTCACTGAGCAGCGTACATC 4408
QY 4321 CCGCTACTGTCAACCGGATCTATTTCTGTTGCAAAAGATGAGTGAATCAATTCAT 4380
DB 4409 CCGCTGCTGTCAACGAGTGTTCAGCCGCGAAGAGATAGGCTGACAGCAATCCCTCAAC 4468
QY 4381 CACCTGTCTACTGCTTTCGACACTACGAGTCCGATGTCACCATATATTTGCTGATTA 4440
DB 4469 CATCTATTTCACAGCAATGAGACGACACGCGACCTGACGTCATCTACTGACAGACAAA 4528
QY 4441 CAATGGAGACCAAGATATGAGGCCATTCAACGCAAGAAAGGCTGAAATTTGAT 4500
DB 4529 AGTTGGAGAGAAATCCAGAAAGCCATGACATGAGAGAGCGCTGTGAGTGTCTCAT 4588
QY 4501 GATGACAAAGCAAGTACAGATTTGCTGAGGGTCCAGCCCAACAGCTTTTGGCAGGC 4560
DB 4589 GATGACGTGAGCTGACACAGACTTGTGTGAGATGACCCGAGACAGACCTGTGGGT 4648
QY 4561 AGACAGGTTATCTCCGTCATAGAGGCAAGTTGATTCATCTGTAAGTACAGATTC 4620
DB 4649 CGTAAGGCTACAGTACAGTGAAGGGTCCGTGATCTGTACTTTGAAGGTACGAAATTC 4708
QY 4621 CATGACCCGCAAGAGCATTTGCCGAATCCATGCAATGTGGCCCAACAAATCTGAGGCT 4680
DB 4709 AACCAAGGCTGCTATGATATGCAAGATGACTGACGTTGTGGCCACAGCTGCAAGAGCA 4768
QY 4681 AATGACGATTTGCTGTACATCTGGGGAGAGATGTCAAGATCCGCTCCAAATGC 4740
DB 4769 AACGAAAGATATGCTTATAGCGCTGGCCGAACAAATGACAAATGATCCAAATGT 4828
QY 4741 CCAGTAGAGAGTACAGAGCGTCTGCTCACCTCAACACTTCATGCTGTGTAATTC 4800
DB 4829 CCGGTAACGATTCGATTCATCAACACTCCCAAGAGAGTGCCTGTGTCGGCGCTAC 4888
QY 4801 GCTATGACGCTGAGCGCGTATACAGTTGGCTCTGCGAAGAAAGAACTGCGCTA 4860
DB 4889 GCAATGACAGCAAGATGCGCCGCTTGTAGTCAACCAAGTTAAAGATGTGTGTT 4948
QY 4861 TGCTCATCATCTCGTTGCCGAGTACAGATCAACAGCGGTGCAAGACTCAAGTGACGC 4920
DB 4949 TGCTCATTTTCCCTCCGAAATACATGATGATGGGGTGCAGAAAGTAAAGTGCAG 5008
QY 4921 AAACAGTCTGTTTTCAGGCGTCTACACCGGCTGTACACCCGAGAAATGACGCGAA 4980
DB 5009 AAGGTTCTCTGTTGACCCGACGCTACTTCAAGTGTAAAGTCCGCGAAGATATGCCCA 5068

```

RESULT 13
ABN86690
ID ABN86690 standard; DNA; 13599 BP.

```

XX AC ABN86690;
XX DT 05-NOV-2002 (first entry)
XX DE Nucleotide sequence of vector pSCA1-E7-Hsp70.
XX KM Major histocompatibility complex; MHC; antigen presenting cell; APC;
XX KM antigen; cytosolic; vintide; gene therapy; CD8; vaccine; therapeutic;
XX KM cancer; viral infection; HPV; E7; heat shock protein 70; Hsp70; da.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT misc_feature 7405..9509
XX FT /tag= a
XX FT /note= "E7-Hsp70 fusion sequence"
XX PN MO20261113-A2.
XX PD 08-AUG-2002.
XX PE 01-FEB-2002; 2002WC-US02598.
XX PR 01-FEB-2001; 2001US-265334P.
XX PA (UYJO ) UNIV JOHNS HOPKINS.
XX PI Wu T, Hung C;
XX DR WPI; 2002-619261/66.
XX PT Nucleic acid molecule encoding a fusion polypeptide that promotes
XX PT processing via the Major Histocompatibility Complex class I pathway
XX PT and/or promotes activity of an antigen presenting cell, useful as
XX PT vaccine for cancer and viral infections -
XX PS Claim 37; Page 34-37; 127pp; English.
XX CC The invention relates to a new nucleic acid molecule (I) encoding a first
XX CC fusion polypeptide useful as a vaccine composition. (I) comprises a first
XX CC nucleic acid sequence encoding a first polypeptide or peptide that
XX CC promotes processing via the Major Histocompatibility Complex (MHC) class
XX CC I pathway (MHC-I-pp) and/or promotes development or activity of an
XX CC antigen presenting cell (APC). The second nucleic acid sequence is linked
XX CC in frame to the first nucleic acid sequence or to a linker nucleic acid
XX CC sequence and encodes an antigenic polypeptide or peptide. The methods and
XX CC compositions of the present invention are useful as therapeutic vaccine
XX CC for cancer and for major viral infections, such as hepatitis and cervical
XX CC cancer, that cause morbidity and mortality. They can also be used in
XX CC treating animal diseases, such as equine herpesvirus, bovine viruses,
XX CC Marek's disease, retroviral and lentiviral diseases and rabies, in the
XX CC veterinary medicine context. The present sequence represents the
XX CC nucleotide sequence of vector pSCA1-E7-Hsp70 comprising the human
XX CC papillomavirus (HPV) E7 antigenic protein and M. tuberculosis heat
XX CC shock protein 70 (Hsp70) fusion sequence.
XX SQ Sequence 13599 BP; 3582 A; 3516 C; 3647 G; 2854 T; 0 other;
QY Query Match 31.2%; Score 1562.4; DB 24; Length 13599;
DB Best Local Similarity 58.3%; Pred. No. 0;
QY Matches 2903; Conservative 0; Mismatches 2041; Indels 36; Gaps 8;
DB 31 AGAATTCACTGATGATGATGCTGACAGCCCGTATGTCAGTGTGTACAGCGGACGTTT 90
QY 95 AAATGATATGTTGATTTAGAGGCTGACAGCCCATTCATCAAGTCTTTCAGAAAGCAATTT 154
DB 91 CCACAATTGATGATGCAAGAGGAGGCTGACGATGCAATGCAATGCAATGCAAGAGCG 150
QY 155 CCGTGTTCAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214
DB 151 TTTTGCATGTGGCAAAAGCTCATTTGAGAGCGAAGTGCAGCCGGAACCAATATCTTG 210

```

Db 215 TTTTCGACCTGGCTACCAATTGATCGAGCAGAGACTGACAAAGACACACTCATCTTG 274
Qy 211 GACATTGGAAGTGGCCCGCTCAGACATGCACTTCGCAATCACCGCATCATTTGATTCGC 270
Db 275 GATATCGGAGTGGCCCTTCCAGAGAAATGATGTCTACGACAAATACCACTGCGTATGC 334
Qy 271 CCTATGATTAAGCGCTGAAGACCAGACATCAACCGTATGAGAAAGACT-----T 324
Db 335 CCTATGCCAGCGCAGAAAGACCCCGAAAGGCTCGATAGCTACGCAAAAGAACTGGCAGCG 394
Qy 325 AAGAAAAGTGAATTACGACAGAAACATAGCTCTAAGCCGCGAGACTGCTGGAATC 384
Db 395 GCCTCCGGAGAGTCTGATAGAGATCGCAGAAAATCAACCGACCTGACAGCCGTC 454
Qy 385 ATGTCAACCCAGACGACAGACTCCATCTGTGTATGACACAGAGCCGCGTATAG 444
Db 455 ATGGCTAGCCAGAGCTGAATCTCTTACCTTTGCTGATACAGACGTCACGATGCT 514
Qy 445 TACTTTGAGAGTGAAGATATACAGATGTGTACGAGTCATGCAACGACATCAATC 504
Db 515 ACGGACCGCAAGTGGCCGATATACAGAGCGTGTATGCTGATCATGCAACATCGCTG 574
Qy 505 TACCAACGAGCGCTTAAGAGATTAAGCAATTTACTGATAGCTTTGACAGACCTT 564
Db 575 TACCATCAGGCGATGAAAGGTGTCAAGACGCGTATTTGATTTGGATTGACACACCCG 634
Qy 565 TTTATGTACAAAACATGGCAGGTTCTACCTACTTTCACACGAACTGGGCTGACAG 624
Db 635 TTTATGTTTGAACGCGCTACGAGCGCGTATTCACACCTACGCAAACTGGGCGACAG 694
Qy 625 AGAGTATTTGAAGACGTAACATTTGCGCTCGTAACTCAGATCTTCAGAGAGCAGGCTT 684
Db 695 CAGGTATACAGGCGCAGAGACATAGGACTGTGTGAGACATCCTTACGTAGAGGAAGACTC 754
Qy 685 GGAACACTCTCATCTTATGAGAAAGAGGCTCAACCTAATTAATTAATTAATTTCTG 744
Db 755 GCGAAACGTCTCATCTCCGCAAGAGCAATTTGAACCTTGGACACAGTCACTGTTCTG 814
Qy 745 GTTGTTCAACAATCTACAGAAATAGATGATCTGTACGTAGAGTGGCGATTTCCAAAC 804
Db 815 GTAGATCTTACATTTATACCTGAGACGAAAGCTATGAGAGCTGGGACTTACCTTCC 874
Qy 805 GTGTTCACTTGAAGAAAGTCTTAATCTCAAGGTATGATGTGGAGCACTTGTACAGTGT 864
Db 875 GTATTCACCTGAAGGTAAACAATCTTTACTGTAGGTGAGATCACTGATCATGCT 934
Qy 865 GAAAGGTACGTCAATAAAGATTAAGATCAGCCAGAGCTATACGTTAAAGTTAGAAC 924
Db 935 GAAAGGTACGTATGAAGAAATCACTATGTGCTCCGCGCTGTACGTTAAACGCTAGAG 994
Qy 925 TTGGGCTCACAAATGCAATCGAGAGGTTCTTGAATGTGAAGTCAAGATACGCTGGC 984
Db 995 TACGCGTGAAGTATCAGCGAGGAGATTTCTTATGTGCAAGACACAGACCTGTCAA 1054
Qy 985 GCGCAGAGGTTTCTTTGCTGTGTATGATGATACAGCACACTTTGCGATCAGATG 1044
Db 1055 GGAAGAAAGTCTATTCCTCTGTATGCACTACGCTCCCTCAACATCTGTGATCAATG 1114
Qy 1045 ACAGGATTTCTGCAACTGACGTTAGTGTATGATGACGCAAAAATCTATTTGGTGGCTC 1104
Db 1115 ACTGCACTATAGCAGACGACGCTCACACCGAGAGACGACAGAAATTTGATGGATTTG 1174
Qy 1105 AACCAAGATTTGTGTCAATGTGTAGAGCGCAAAATATCTTAACATATGCAAACTAT 1164
Db 1175 AATCAGAGATATGTGTGAACGAAACACAGCAAACTATTAACATATGAAAGACTAT 1234
Qy 1165 CTATTACAGTGTGCGCCAGGCGTTTTCAGAGTGGCGCGTGAACATCTGTCGACCTTG 1224
Db 1235 CTGCTTCGATTTGTGGCGCTGCACTTTAGCAAGTGGCGAGGAAATCAAGCACAACCTT 1294
Qy 1225 GACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTTACTATGGGCTGCTGTGGCT 1284
Db 1295 GATGATGAAAAAATCTGTGGGTGTCGAGAGAGGTCACTTACTTGTGCTGTGTGGCA 1354

Qy 1285 TTCAAGACCCAGAAATATCATTCATCTACAGAAAGCTTGGTACGCAAACTATAGAAA 1344
Db 1355 TTTAAAGAGAGAAATATCACAACCTATCAAGAAACCAAGACCCACAAATATAGTAAG 1414
Qy 1345 GTACCTGCGCTTTTGAATCTTGTGTATTCACGCTTACCAAGCCAGCGGCTGATATG 1404
Db 1415 GTGCTTCAAGATTTAACTCGTGTGTATCCGAGCCTATGGTTACAGGCTCCCAATC 1474
Qy 1405 GCGTTCCGCGTAGCTCAAGCTGTGCTTGAACCACTGTCAAAACCGCACCGCTATT 1464
Db 1475 CGAGTCATACAGGATTAATGCTTTTGGCCAAAGAAAGACCAACG---GAGATTAATA 1531
Qy 1465 ACAATGGCGATGTGAGACATCTGCGGCTTACAGCAAGAGCTGAAGAAATGGCTGCA 1524
Db 1532 CTGTGTTCTGACGCGTGTCTACCGCAGGAGTGTGAACAAAGAGAGAGAGTTGGAG 1591
Qy 1525 GCGGAAGATCAGAAAGCCCTGCAACCTTGTCTCCCTGAATAGAAAAAGAGCCGTA 1584
Db 1592 GCGGAGTGACTAGAGAGCCTTACACCCCTGTCCTCCATCGCGCGGAGAGCGGGA 1651
Qy 1585 GAGCGAAGATGACCTTATATGCAAG-----GAGAGACGATAGCTGAGACA 1638
Db 1652 GTCTGACGCTGACGTTGAAGAACTAGAGTATCACGACAGTGCAGGGGTGCTGGAACA 1711
Qy 1639 CGACGAGACACATCAGGGTGAACAAGTTACCGACGAGAAAGAAATTTGGGCTTACGCT 1698
Db 1712 CTGCGACGCGTTGAAGTCAACGACACGCGCAACGACGTAATCTAGAAATTTACGTA 1771
Qy 1699 ATACTTTCACCCAGGCGGTATTTGAATGTAGAAAACTGGCGGTATTCACCCATTTGCG 1758
Db 1772 GTTGTCTCCCGCAGACCGTGTCTCAAGGCTTCGAAGTTGGCCCGCTGCACTCTTACCA 1831
Qy 1759 GACAAGTATGTATATATCTCAAAAGGTAGGCGAGGAGATTAAGTGAAGCATAC 1818
Db 1832 GAGCAGGTGAATAATTAATTAACAATTAACGAGGCGCGCGGTATACAGATGACGAGAT 1891
Qy 1819 CACGTTAAGTCAATTTATCCAGAAAGGAGCGGCGCTGTTCAAGACTTCAGGATTTG 1878
Db 1892 GAGGCAAGGTCATCTACATGTGTATGATCGGCTATTCGAGTCCCTGATTTCAAGCTTTG 1951
Qy 1879 AGTGAACGCTACGATCGTTTTCACAGAGAGAGTTCGTAAACAGATACCTGCACAC 1938
Db 1952 AGCGAGAGCGCACTATGTGTATCAAGAAAGGAGTTCGTCAACAGAACTATACAT 2011
Qy 1939 ATCGCAATCAACGAGAGCGCTTAACATGACGAGATCTATTAAGCTGTAAAGCT 1998
Db 2012 ATTGCCGTTCAACGACCGCTGCTGAACCGACGAGGAGAACTACGAAAGATCAGACT 2071
Qy 1999 CAGACACAGACTAGAAATAGTCTTGTGAATTTGACGACGAAAGTGTATTAAGCGAAA 2058
Db 2072 GAAGAACTGACCGCAAGTATGCTGTGACGATAGTAATAAATGCTGTGCTCAAGAGAG 2131
Qy 2059 GACGAGGTCCCTTGTGCTTAAACGAGTATCTGTATATCCACATTTACAGAGTTTGC 2118
Db 2132 GAAGGTGCGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2191
Qy 2119 TACGAGTCTCAAGACAGACAGACACTTCAAAAGTCCCAACCATCGAGCTAT 2178
Db 2192 TACGAAAGGCTGAAGATCAGCGCTCGGCAACATATAGATACAGTATAGTGTCTTT 2251
Qy 2179 GAGAGTCCAGGTTCAAGTAACTGTGAATCTCAAAAGCGCTGTGATCTTAAGAAAGTCTG 2238
Db 2252 GGGGTTCCGGATCAGGCAAGTCTGTATTTAAGAACTGTGTGACCAAAACAGATCTG 2311
Qy 2239 GTTGTAGTGTGGAAGAAAGAACTGCGCAGAAATCATCAGGAGTGAAGAGATGAGA 2298
Db 2312 GTCAACAGCGGCAAGAGAGAACTGTCCAGAAATATGTTAAGACGTGAAGAGACCGC 2371
Qy 2299 GGTATGATGTTGCTCTAGAGCTGTGATTAAGTGTCTTTAATTTGGGTTAAGACCCC 2358
Db 2372 GCGAAGGAGCAAGATAGGAAACAGATGATCTCTGTAAACGAGGTGTCTGTGCTC 2431

Qy 2359 GTTAACCTGTACATGATGAGCATTTGCTGCCATGCGAGGACGCTGCGACTG 2418
Db 2432 GTGGAACATCTATATGTGAGAGAGGCTTTGCTAGCCATTCGGTATCTCTGCGCTTA 2491
Qy 2419 ATTGCCATCGTCAACCT--AAGAAAGTGTATTTGTGCGGGGACCCAAAACAATGCGGC 2475
Db 2492 ATTGCTCTTGTAAACCTCGAGCAAAAGTGTATTATCGGAGACCCCAACAAATGCGGA 2551
Qy 2476 TTCTTTAACTGATGTGCTGAAAGTACATTTTAAACCTGACATATGACATGAAAGTGTAC 2535
Db 2552 TTCTTCAATATGATGACGCTTAAGGTGAACCTTCAACCAACATCTGACATGAAGTATGT 2611
Qy 2536 CATAAAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATGCTGCCACTCTTC 2595
Db 2612 CATAAAGTATATCAAGACGTTGCAAGGCTGACAGGCGCATCGTGTCTACGTTGCAC 2671
Qy 2596 TAGACAAGCGCAATGAAGCGTTAAACCATGTGCTGTATTAATCATATGATACACA 2655
Db 2672 TAGGAGGCAAGATGCGCACAGCAACCCGTCAACAAACCATATATCATAGACACACA 2731
Qy 2656 GGGACACAAAGCCGCAACAAAGATGATGATTTCTAAACCTGTTCAAGGATGGGTAAA 2715
Db 2732 GGACAGACCAAGCCCAAGCCAGAGACATCGTTTAAACATGCTTCCGAGGCTGGCAAG 2791
Qy 2716 CAGCTACAGATTGACTACAAATAATCAAGAAATCAATGCTGCGGTGACATCGCAAGACTT 2775
Db 2792 CAGCTGACGTTGATCTACCGTGACACGAACATCATAGACAGACGATCTCAGGGCTC 2851
Qy 2776 ACGCGAAAGCGTTTATGCTGTGACAGTAAAGTCAACGAAATCCACTCTACTCGAG 2835
Db 2852 ACCCGCAAGGGGTATACGCCGTAAAGCAGAGGATGAAATCCCTGTATGCCCCCT 2911
Qy 2836 ACTCTAGACGCTGAACGCTTACTTACAGCAACAGAAACCAATGCTGTGAAGACG 2895
Db 2912 GCGTGGAGCACGTAATGTACTGCTGACGCACTGAGATPAGCTGTGTGAAGAACG 2971
Qy 2896 CTAACTGGTGAATCCCTGGATTAAGACACTTACAGTAAATATCCCGGGGATTTTACAGGCT 2955
Db 2972 CTGGCCCGGCAATCCCTGGATTAAGTCTTATCAAAACATTTCCACAGGTTACTTTACGGCC 3031
Qy 2956 TCATTTGACGACTGCGACGCGGAACAGACGCCATTAATGACGCGTTCTTGATPAGCGC 3015
Db 3032 ACATTTGAAAGATGGCAAGAAACAACGAAATTAATGAAGTGAATGAAGACACGGCT 3091
Qy 3016 CAGACAGCTGATGTGTTCCAAATAATPAGTGAACGCTGCTGGGGGAAAGCTTTAGAGCA 3075
Db 3092 GCGGCTGTGAGCGGTTCCAAACAAACGAAACGTTGTGGGAAAGCGCTGTGCTT 3151
Qy 3076 GTCTTGGCCACGCGCAACATTTGTCTGACGAGACAGACAGTGGAGAC--GTTGACCCA 3132
Db 3152 GTCTTGGACATGCGCGGAATCAGATTGACACAGAGAGTGGACACCATATATTACAGCA 3211
Qy 3133 TTCAAGCATGACAGAGCTACTCACTGAAAATGCACTGAACCTTCTTTTGCACAGGTTT 3192
Db 3212 TTTAAGGGGACAAAGCTTACTCTCAAGTGTGCGCTTGAATGAATTTGCACCAAGTAC 3271
Qy 3193 TTTGAGAGAGACTGAGACAGTGGGTTATTTTCCGCTCTACCGTGCATTTACTAACG 3252
Db 3272 TATGAGTTGACCTGGAAGTGGCTGTTTTTCTGCCGAAGGTGCTCTGTATTATACAG 3331
Qy 3253 GATAGACACTGGGATTAATCTGCCAGAGGAACATGTATGGGCTTAATAGAGAGTAGCA 3312
Db 3332 AACAACCACTGGGATTAACAGACTGTGTGAAAGATGTATGATTTCAATGCCGCAACACT 3391
Qy 3313 AAGAGTTGTACGCGCATATCTGCTGATCACAAAAGCGTTGACACAGGAGGTTACT 3372
Db 3392 GCCAGGCTGAAGCTAGACATACCTTCTGAAGGGGAGTGGACATACGGGCAACAGGCA 3451
Qy 3373 GATATTAAGATTAATACATCAAGACTACTCTCAACAAATTAATATGTGTCTCAATTAAT 3432
Db 3452 GTTATTCGAGAAAGAAATCAACCGCTTTCTGTGTGACAAATGTATTAATCTATCAAC 3511
Qy 3433 CGCGGTTGCCCACTGTTGATGTTGACCAACAAAGGAGGGAACAACTGATCACAGC 3492

Db 3512 CGAGGCTGCGGACCGCCCTGTGGTGAATACAAAGCGTTAAAGCAGATAGGTTAG 3571
Qy 3493 GAATTCCTATCTAATGAAAGGCAATCTGTGTGTGATTCGCGCATCTATATACGATT 3552
Db 3572 TGGCTGTCTAATTAAGTAAGAGGTTACCAAGTCTCTGTGTGTGATGATCAACCTGGCT 3611
Qy 3553 CAGGGAAGAAAGTAAAGTGCATGGGTCCATGCGCCCATATAC-----CATCAGG 3603
Db 3632 TTGCTTGAAGCAGGATCACTTGGTTGTCAACGCTGAATGTCAAGCGCCGATAGGTC 3691
Qy 3604 TGTGATCTGATTTGGGAATACCTAAGCATGTGCTGAATATGACATTAATCTTTGCAAT 3663
Db 3692 TAGACTTAAGTTTAGACTGCGCGGTGAGCGCGGAGGTTGCACTTGTGTTGTAAAC 3751
Qy 3664 GTTAGACCCCGTACAGAAACATCACTACCAACATGTCGAGATACCGTATTCACAC 3723
Db 3752 ATTCAACGGAAATTCAAATTCACACTACACAGTGTGTCAACCGCATGAACTG 3811
Qy 3724 AGCATGCTAAGTGTAAAGGCTGTCCACCACTGAAACCTGGCGGAACATGTGGCTATA 3783
Db 3812 CAGATGCTTGGGAGATGCGCTACACATGCTTAAACCTCGCGCATCT--TGATAGA 3868
Qy 3784 GGGTATGGGCTTGTGATTCGCGCAACCGAATATCATCTGCGGTGCAAGCTCATTT 3843
Db 3869 GCTTAGGATACGCGATTAATATCAGGAAGCGGTGTTCTTCTTAACAGAAAGTTT 3928
Qy 3844 AGGTTTACCCGTGTCTGTACGCTTAAGACATGCGCGAAATATCTAGGTTCTTCTGTG 3903
Db 3929 TCGTGTCAAGAGTGTGCGCCCGGATTTGTCAACGCAATTCAGAAAGTGTCTGTG 3988
Qy 3904 TTCTTGGCAAGGCAACGGCAACCAACATGACAGGACAGACTCGGTGTAGTCTT 3963
Db 3989 TTCTTCACTTTTGAACAGAAAGAGACCTCTACGCTAACACGATGAATACCAAGCTG 4048
Qy 3964 GACAACATCTATCAAGGTTCAACAGGTACAGGACAGAGAGCTCCAGGTTACAGATG 4023
Db 4049 AGTGGCTGTATGCGGAGAAAGCATGCAACGCGCGGTGTGACCATCTACAGAGTT 4108
Qy 4024 ATCAGAGTGAATTTAGCAAGAGCGCTGACCAAGCTATGTTAATGTCTAATAGCAAA 4083
Db 4109 AAGAGACCAACATTAAGCAAGTGTCAACAGAGCGGCTGTGTAAACGAGTTAACGCCGT 4168
Qy 4084 GGTCAACGAGTTCCGGAAGTGTGCGGTGACCTGATCCGAATATGGCGGCTGCTTTGAT 4143
Db 4169 GGAACGTAAAGGATGCGTATGATGAGGCGCTGTGCGGAAGAAATGGCGCTTAAG 4228
Qy 4144 AGACAGCAATAGCTGTGCGGACGCTTGAATTGTGAAGACGAAACG--CTCATGATA 4200
Db 4229 GAGACACACACAGTGGGCAATTAACAGTCAATGTGCGGCTGTACCCGTCATC 4288
Qy 4201 CATGCTGTAGAACCAATTTTTTCTAAGATGCGGAAACCGGAGGCGCACTTAAGCTGCA 4260
Db 4289 CACGCTGTACGCTTAATTTTCTGTGCAACACTGAAACCGGAAGGGACCCGGAATTTGGCC 4348
Qy 4261 GTCGCTTAAGAGCATAGGCTCATGTCAACGCTGACGGGATTAACAAATAATACGTA 4320
Db 4349 GCTGTCTAACGGGCAAGTGGCGCGCAAGTAAACAGACTGTCACTAGACAGTATGCCATC 4408
Qy 4321 CCGCTACTGTCAACCGGCATCTATTCTGTGTGCAAAAGATGAGTATGATCAATTCAT 4380
Db 4409 CCGCTGTGTCCACAGAGAGTTCACAGGCGGGAAGATATGGCTGACGACATCCCTCAAC 4468
Qy 4381 CACTGTTCACCTGTTTGCACATACAGATGCGGATGTCAACATATATTGCTGTGATAAA 4440
Db 4469 CATCTATTACACCAATGAAGCGCACGAGCTGACCTGACATCTACACAGACAAA 4528
Qy 4441 CAATGGAGACAGGATTAATCGAGCATTCACCGCAAAAGAAAGCGTCAAAATTCGTGAT 4500
Db 4529 AGTTGGAGAAAGAAATTCAGAAAGCATTTGAATATGAGACGCGCTGTGGAGTTGCTCAAT 4588
Qy 4501 GATGACAGCGAGTGAATTAATGATTTGTCAGGCTCACCCAAACAGCTTTTGGCAGGC 4560

Db 4589 GATGACGTGAGCTGACACAGACTTGTGAGAGTGACCCGAGACGACCTGTGGGT 4648
Qy 4561 AGACCGAGTTACTCCGTCATATGAGGCAAGTTGTATTCTATCTGAAAGTACACGATTTC 4620
Db 4649 CGTAGGGCTTACAGTACACCTGACGGTTCGCTGCTACTGTTGAAAGTACGAAATTC 4708
Qy 4621 CATCAGACCGCCAGACATTTGCCGAATTCATGCAATGTGGCCCAACAAATCTGAGCT 4680
Db 4709 AACCGGCTGTATTGATATGACAGAGATACGATGTTGTGGCCAGACTGCAGAGGCA 4768
Qy 4681 AATGACGATTTGTTGTATGACATCTGGGGGAGATATGTCAGATCCGCTCCAAATGC 4740
Db 4769 AACGACGATATGCTTATACCGCTGCGGAAACATGACAAATCAGATCCAAATGT 4828
Qy 4741 CCAGTAGAGAGTCAAGAGCGCTGCTCCACTCAACACTTCGATGCGTGTATATAC 4800
Db 4829 CCGGTGAACGATTCGATTCATCAACCTCCAGAGACGTGCTGCTGTCGCGCTAC 4888
Qy 4801 GCTATGACGGCTGACGGCGTATACAGTTGCGCTCTGCGAAGAAAGAAACAGTTCCCGTA 4860
Db 4889 GCAATGACAGGACGAGATGCGCGCTTAGTACACACCAAGTAAAGCATGTGTT 4948
Qy 4861 TGCTCATCTTCTCTGTTGCCGAATGACAGATCAGAGCGGTGACAGAGCTACAGTGCAC 4920
Db 4949 TGCTCATCTTCTCTGTTGCCGAATGACAGATGAGAGGTGACAGAGGTAAAGTGCAG 5008
Qy 4921 AACACGCTCTGTTTTCAGGCGGTGTACACCGCGTGTACACCCAGAGAGTACGCGGA 4980
Db 5009 AAGGTTCTCTGTTTCGACCCGACGAGTACCTTAGTGTAGTCCGCGAAGATATCCGCA 5068

RESULT 14
AAQ26021
ID AAQ26021 standard; DNA; 11517 BP.
XX AAQ26021;
AC 25-MAR-2003 (updated)
DT 05-JAN-1993 (first entry)
XX
DE pSP6-SFV4 RNA transcript as DNA.
XX
KW Semliki forest virus; SFV; SP6 expression vector; RNA polymerase; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 87..7379
FT /*tag= a
FT /label= Non-structural_polyprotein
FT CDS 7421..11182
FT /*tag= b
FT /label= Structural_polyprotein
FT
FT
PN MO9210578-A1.
XX
PD 25-JUN-1992.
XX
PF 12-DEC-1991; 91WO-SE00855.
XX
PR 13-DEC-1990; 90SR-0003978.
XX
PA (BIOP-) BIOPTION AB.
P1 Garoff H, Liljestrom P;
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
P12
P13
P14
P15
P16
P17
P18
P19
P20
P21
P22
P23
P24
P25
P26
P27
P28
P29
P30
P31
P32
P33
P34
P35
P36
P37
P38
P39
P40
P41
P42
P43
P44
P45
P46
P47
P48
P49
P50
P51
P52
P53
P54
P55
P56
P57
P58
P59
P60
P61
P62
P63
P64
P65
P66
P67
P68
P69
P70
P71
P72
P73
P74
P75
P76
P77
P78
P79
P80
P81
P82
P83
P84
P85
P86
P87
P88
P89
P90
P91
P92
P93
P94
P95
P96
P97
P98
P99
P100
P101
P102
P103
P104
P105
P106
P107
P108
P109
P110
P111
P112
P113
P114
P115
P116
P117
P118
P119
P120
P121
P122
P123
P124
P125
P126
P127
P128
P129
P130
P131
P132
P133
P134
P135
P136
P137
P138
P139
P140
P141
P142
P143
P144
P145
P146
P147
P148
P149
P150
P151
P152
P153
P154
P155
P156
P157
P158
P159
P160
P161
P162
P163
P164
P165
P166
P167
P168
P169
P170
P171
P172
P173
P174
P175
P176
P177
P178
P179
P180
P181
P182
P183
P184
P185
P186
P187
P188
P189
P190
P191
P192
P193
P194
P195
P196
P197
P198
P199
P200
P201
P202
P203
P204
P205
P206
P207
P208
P209
P210
P211
P212
P213
P214
P215
P216
P217
P218
P219
P220
P221
P222
P223
P224
P225
P226
P227
P228
P229
P230
P231
P232
P233
P234
P235
P236
P237
P238
P239
P240
P241
P242
P243
P244
P245
P246
P247
P248
P249
P250
P251
P252
P253
P254
P255
P256
P257
P258
P259
P260
P261
P262
P263
P264
P265
P266
P267
P268
P269
P270
P271
P272
P273
P274
P275
P276
P277
P278
P279
P280
P281
P282
P283
P284
P285
P286
P287
P288
P289
P290
P291
P292
P293
P294
P295
P296
P297
P298
P299
P300
P301
P302
P303
P304
P305
P306
P307
P308
P309
P310
P311
P312
P313
P314
P315
P316
P317
P318
P319
P320
P321
P322
P323
P324
P325
P326
P327
P328
P329
P330
P331
P332
P333
P334
P335
P336
P337
P338
P339
P340
P341
P342
P343
P344
P345
P346
P347
P348
P349
P350
P351
P352
P353
P354
P355
P356
P357
P358
P359
P360
P361
P362
P363
P364
P365
P366
P367
P368
P369
P370
P371
P372
P373
P374
P375
P376
P377
P378
P379
P380
P381
P382
P383
P384
P385
P386
P387
P388
P389
P390
P391
P392
P393
P394
P395
P396
P397
P398
P399
P400
P401
P402
P403
P404
P405
P406
P407
P408
P409
P410
P411
P412
P413
P414
P415
P416
P417
P418
P419
P420
P421
P422
P423
P424
P425
P426
P427
P428
P429
P430
P431
P432
P433
P434
P435
P436
P437
P438
P439
P440
P441
P442
P443
P444
P445
P446
P447
P448
P449
P450
P451
P452
P453
P454
P455
P456
P457
P458
P459
P460
P461
P462
P463
P464
P465
P466
P467
P468
P469
P470
P471
P472
P473
P474
P475
P476
P477
P478
P479
P480
P481
P482
P483
P484
P485
P486
P487
P488
P489
P490
P491
P492
P493
P494
P495
P496
P497
P498
P499
P500
P501
P502
P503
P504
P505
P506
P507
P508
P509
P510
P511
P512
P513
P514
P515
P516
P517
P518
P519
P520
P521
P522
P523
P524
P525
P526
P527
P528
P529
P530
P531
P532
P533
P534
P535
P536
P537
P538
P539
P540
P541
P542
P543
P544
P545
P546
P547
P548
P549
P550
P551
P552
P553
P554
P555
P556
P557
P558
P559
P560
P561
P562
P563
P564
P565
P566
P567
P568
P569
P570
P571
P572
P573
P574
P575
P576
P577
P578
P579
P580
P581
P582
P583
P584
P585
P586
P587
P588
P589
P590
P591
P592
P593
P594
P595
P596
P597
P598
P599
P600
P601
P602
P603
P604
P605
P606
P607
P608
P609
P610
P611
P612
P613
P614
P615
P616
P617
P618
P619
P620
P621
P622
P623
P624
P625
P626
P627
P628
P629
P630
P631
P632
P633
P634
P635
P636
P637
P638
P639
P640
P641
P642
P643
P644
P645
P646
P647
P648
P649
P650
P651
P652
P653
P654
P655
P656
P657
P658
P659
P660
P661
P662
P663
P664
P665
P666
P667
P668
P669
P670
P671
P672
P673
P674
P675
P676
P677
P678
P679
P680
P681
P682
P683
P684
P685
P686
P687
P688
P689
P690
P691
P692
P693
P694
P695
P696
P697
P698
P699
P700
P701
P702
P703
P704
P705
P706
P707
P708
P709
P710
P711
P712
P713
P714
P715
P716
P717
P718
P719
P720
P721
P722
P723
P724
P725
P726
P727
P728
P729
P730
P731
P732
P733
P734
P735
P736
P737
P738
P739
P740
P741
P742
P743
P744
P745
P746
P747
P748
P749
P750
P751
P752
P753
P754
P755
P756
P757
P758
P759
P760
P761
P762
P763
P764
P765
P766
P767
P768
P769
P770
P771
P772
P773
P774
P775
P776
P777
P778
P779
P780
P781
P782
P783
P784
P785
P786
P787
P788
P789
P790
P791
P792
P793
P794
P795
P796
P797
P798
P799
P800
P801
P802
P803
P804
P805
P806
P807
P808
P809
P810
P811
P812
P813
P814
P815
P816
P817
P818
P819
P820
P821
P822
P823
P824
P825
P826
P827
P828
P829
P830
P831
P832
P833
P834
P835
P836
P837
P838
P839
P840
P841
P842
P843
P844
P845
P846
P847
P848
P849
P850
P851
P852
P853
P854
P855
P856
P857
P858
P859
P860
P861
P862
P863
P864
P865
P866
P867
P868
P869
P870
P871
P872
P873
P874
P875
P876
P877
P878
P879
P880
P881
P882
P883
P884
P885
P886
P887
P888
P889
P890
P891
P892
P893
P894
P895
P896
P897
P898
P899
P900
P901
P902
P903
P904
P905
P906
P907
P908
P909
P910
P911
P912
P913
P914
P915
P916
P917
P918
P919
P920
P921
P922
P923
P924
P925
P926
P927
P928
P929
P930
P931
P932
P933
P934
P935
P936
P937
P938
P939
P940
P941
P942
P943
P944
P945
P946
P947
P948
P949
P950
P951
P952
P953
P954
P955
P956
P957
P958
P959
P960
P961
P962
P963
P964
P965
P966
P967
P968
P969
P970
P971
P972
P973
P974
P975
P976
P977
P978
P979
P980
P981
P982
P983
P984
P985
P986
P987
P988
P989
P990
P991
P992
P993
P994
P995
P996
P997
P998
P999
P1000
P1001
P1002
P1003
P1004
P1005
P1006
P1007
P1008
P1009
P1010
P1011
P1012
P1013
P1014
P1015
P1016
P1017
P1018
P1019
P1020
P1021
P1022
P1023
P1024
P1025
P1026
P1027
P1028
P1029
P1030
P1031
P1032
P1033
P1034
P1035
P1036
P1037
P1038
P1039
P1040
P1041
P1042
P1043
P1044
P1045
P1046
P1047
P1048
P1049
P1050
P1051
P1052
P1053
P1054
P1055
P1056
P1057
P1058
P1059
P1060
P1061
P1062
P1063
P1064
P1065
P1066
P1067
P1068
P1069
P1070
P1071
P1072
P1073
P1074
P1075
P1076
P1077
P1078
P1079
P1080
P1081
P1082
P1083
P1084
P1085
P1086
P1087
P1088
P1089
P1090
P1091
P1092
P1093
P1094
P1095
P1096
P1097
P1098
P1099
P1100
P1101
P1102
P1103
P1104
P1105
P1106
P1107
P1108
P1109
P1110
P1111
P1112
P1113
P1114
P1115
P1116
P1117
P1118
P1119
P1120
P1121
P1122
P1123
P1124
P1125
P1126
P1127
P1128
P1129
P1130
P1131
P1132
P1133
P1134
P1135
P1136
P1137
P1138
P1139
P1140
P1141
P1142
P1143
P1144
P1145
P1146
P1147
P1148
P1149
P1150
P1151
P1152
P1153
P1154
P1155
P1156
P1157
P1158
P1159
P1160
P1161
P1162
P1163
P1164
P1165
P1166
P1167
P1168
P1169
P1170
P1171
P1172
P1173
P1174
P1175
P1176
P1177
P1178
P1179
P1180
P1181
P1182
P1183
P1184
P1185
P1186
P1187
P1188
P1189
P1190
P1191
P1192
P1193
P1194
P1195
P1196
P1197
P1198
P1199
P1200
P1201
P1202
P1203
P1204
P1205
P1206
P1207
P1208
P1209
P1210
P1211
P1212
P1213
P1214
P1215
P1216
P1217
P1218
P1219
P1220
P1221
P1222
P1223
P1224
P1225
P1226
P1227
P1228
P1229
P1230
P1231
P1232
P1233
P1234
P1235
P1236
P1237
P1238
P1239
P1240
P1241
P1242
P1243
P1244
P1245
P1246
P1247
P1248
P1249
P1250
P1251
P1252
P1253
P1254
P1255
P1256
P1257
P1258
P1259
P1260
P1261
P1262
P1263
P1264
P1265
P1266
P1267
P1268
P1269
P1270
P1271
P1272
P1273
P1274
P1275
P1276
P1277
P1278
P1279
P1280
P1281
P1282
P1283
P1284
P1285
P1286
P1287
P1288
P1289
P1290
P1291
P1292
P1293
P1294
P1295
P1296
P1297
P1298
P1299
P1300
P1301
P1302
P1303
P1304
P1305
P1306
P1307
P1308
P1309
P1310
P1311
P1312
P1313
P1314
P1315
P1316
P1317
P1318
P1319
P1320
P1321
P1322
P1323
P1324
P1325
P1326
P1327
P1328
P1329
P1330
P1331
P1332
P1333
P1334
P1335
P1336
P1337
P1338
P1339
P1340
P1341
P1342
P1343
P1344
P1345
P1346
P1347
P1348
P1349
P1350
P1351
P1352
P1353
P1354
P1355
P1356
P1357
P1358
P1359
P1360
P1361
P1362
P1363
P1364
P1365
P1366
P1367
P1368
P1369
P1370
P1371
P1372
P1373
P1374
P1375
P1376
P1377
P1378
P1379
P1380
P1381
P1382
P1383
P1384
P1385
P1386
P1387
P1388
P1389
P1390
P1391
P1392
P1393
P1394
P1395
P1396
P1397
P1398
P1399
P1400
P1401
P1402
P1403
P1404
P1405
P1406
P1407
P1408
P1409
P1410
P1411
P1412
P1413
P1414
P1415
P1416
P1417
P1418
P1419
P1420
P1421
P1422
P1423
P1424
P1425
P1426
P1427
P1428
P1429
P1430
P1431
P1432
P1433
P1434
P1435
P1436
P1437
P1438
P1439
P1440
P1441
P1442
P1443
P1444
P1445
P1446
P1447
P1448
P1449
P1450
P1451
P1452
P1453
P1454
P1455
P1456
P1457
P1458
P1459
P1460
P1461
P1462
P1463
P1464
P1465
P1466
P1467
P1468
P1469
P1470
P1471
P1472
P1473
P1474
P1475
P1476
P1477
P1478
P1479
P1480
P1481
P1482
P1483
P1484
P1485
P1486
P1487
P1488
P1489
P1490
P1491
P1492
P1493
P1494
P1495
P1496
P1497
P1498
P1499
P1500
P1501
P1502
P1503
P1504
P1505
P1506
P1507
P1508
P1509
P1510
P1511
P1512
P1513
P1514
P1515
P1516
P1517
P1518
P1519
P1520
P1521
P1522
P1523
P1524
P1525
P1526
P1527
P1528
P1529
P1530
P1531
P1532
P1533
P1534
P1535
P1536
P1537
P1538
P1539
P1540
P1541
P1542
P1543
P1544
P1545
P1546
P1547
P1548
P1549
P1550
P1551
P1552
P1553
P1554
P1555
P1556
P1557
P1558
P1559
P1560
P1561
P1562
P1563
P1564
P1565
P1566
P1567
P1568
P1569
P1570
P1571
P1572
P1573
P1574
P1575
P1576
P1577
P1578
P1579
P1580
P1581
P1582
P1583
P1584
P1585
P1586
P1587
P1588
P1589
P1590
P1591
P1592
P1593
P1594
P1595
P1596
P1597
P1598
P1599
P1600
P1601
P1602
P1603
P1604
P1605
P1606
P1607
P1608
P1609
P1610
P1611
P1612
P1613
P1614
P1615
P1616
P1617
P1618
P1619
P1620
P1621
P1622
P1623
P1624
P1625
P1626
P1627
P1628
P1629
P1630
P1631
P1632
P1633
P1634
P1635
P1636
P1637
P1638
P1639
P1640
P1641
P1642
P1643
P1644
P1645
P1646
P1647
P1648
P1649
P1650
P1651
P1652
P1653
P1654
P1655
P1656
P1657
P1658
P1659
P1660
P1661
P1662
P1663
P1664
P1665
P1666
P1667
P1668
P1669
P1670
P1671
P1672
P1673
P1674
P1675
P1676
P1677
P1678
P1679
P1680
P1681
P1682
P1683
P1684
P1685
P1686
P1687
P1688
P1689
P1690
P1691
P1692
P1693
P1694
P1695
P1696
P1697
P1698
P1699
P1700
P1701
P1702
P1703
P1704
P1705
P1706
P1707
P1708
P1709
P1710
P1711
P1712
P1713
P1714
P1715
P1716
P1717
P1718
P1719
P1720
P1721
P1722
P1723
P1724
P1725
P1726
P1727
P1728
P1729
P1730
P1731
P1732
P1733
P1734
P1735
P1736
P1737
P1738
P1739
P1740
P1741
P1742
P1743
P1744
P1745
P1746
P1747
P1748
P1749
P1750
P1751
P1752
P1753
P1754
P1755
P1756
P1757
P1758
P1759
P1760
P1761
P1762
P1

QY 925 TTGGCGTCCAAATGCAATCGGAGGGTTTCTTGAGTGGCAAGATCAAGATACCGTGGC 984
DB 996 TACCGCGTGAAGTATCAACGCGAGGGAATTCCTAGTGTCAAGACACAGACACTGTCAA 1055
QY 985 GCGCAGAGGGTTCTTTTGTCTGTGTATCGATGTACAGCCACACTTTGGATCAATG 1044
DB 1056 GGAAGAAAGATCTATCCCTGTATGACCTACCTCCCTCAACCATCTGTGATCAATG 1115
QY 1045 AACGGGATTCGGCACTGACGTATAGTGTGATAGCGACAAAACATTTGGTGGGCTC 1104
DB 1116 ACTGGCACTACAGGACCGAGCTCACACGGAGACGACAGAAATGTGTATGAGGATG 1175
QY 1105 AACCAAGATTTGTCTCAATGTAGAGCGCAAGAAAATTAACAACAATGACAGATCT 1164
DB 1176 AATAGAGATAGTGTGTGAAGGAAGAACAGGAAACATTAACAGATGAATGACTAT 1235
QY 1165 CTATTAACAGTGTGCGCCAGGCGTTTTCAGGTGGGCGGTGAACATCGTCCGACTTG 1224
DB 1236 CTGCTTCGATTTGTGGCGGTGCAATTAAGCAAGTGGGGAAGAAATCAAGGACAGACTT 1295
QY 1225 GACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTTACTATGGGCTGCTGGGCT 1284
DB 1296 GATGATAAAAAAGCTCTGGGTGTCCGAGAGGTCACTTACTTGTCTGTGTGTGCA 1355
QY 1285 TTCAAGACCCGAGAAATCAATCCATCTACAAAGAGCGGTAGCGAAACATTAAGAA 1344
DB 1356 TTTAAAAAGAGAAAGATGACACCAATGTACAAAGAACAGACCCAGACAAATGTGAG 1415
QY 1345 GTACCTGCGCTCTTTGACTCATTTGTGTATTCACGCTTACCAAGCGGCTCGATATG 1404
DB 1416 GTGCTTCAGAGTTTAATCGTTCGTCACTCCGAGCTATGCTTACAGGCGCTCGCAATC 1475
QY 1405 GCGTTCGCGCGTACGCTCAAGCTGTGTTGAACCACTGTCAACCCGCAACCGGCTAT 1464
DB 1476 CCACTCAAGATCAACGATTAAGATCTTTGGCAAGAAAGCAAGC---GAGAGTTAATA 1532
QY 1465 ACAATGCGCGATGTGAGCATCTGCTGCTTACAGCAAGAGTGAAGAGTGGCTGCA 1524
DB 1533 CTTGTTCTCAACGCGGTGTCAAGCAGGATGTGTAAACAGAGAGAAATGTGGAG 1592
QY 1525 GCGGAGAGATCAGAGAAAGCCCTGCAACCTTGTCTCCCTGAATAGAAAAAGACGCTA 1584
DB 1593 GCCAGCTGACTAGAGAAAGCTTACCAACCCCTCGTCCCATCGCGCGGAGACGGGA 1652
QY 1585 GAGGCAAGATGATCTCATATATGCAAGAG-----GAGAGACAGATAGGTGTGAGCA 1638
DB 1653 GTCTGCAAGTGAAGCTGTTGAAGAACTAGAGATACAGGAGGTGCAAGGGGTGTGAAACA 1712
QY 1639 CCAAGAGACATCAAGGCTGACAAAGTTAACAGGCGAAGAGATTTGGTCTTACGCT 1698
DB 1713 CTTGCGACGCGGTGAAGATCAACGACAGCCGAACGATCTACTAGGAAATTAAGTA 1772
QY 1699 ATACTTTCAACCCAGCGCGTATTAATAGTAAAAATCGCGGTATACCACTTGGCG 1758
DB 1773 GTTCTGTCCCGGACCGGTCTCAAGAGCTCAAGTTGGCCCCGCTCACTCTAGCA 1832
QY 1759 GAAACAATCTGTATGATCTCAAAAGGTAGGGGAGGAAATCAAAAGTGAAGCATATC 1818
DB 1833 GAGAGGTGAATAATTAACATTAACGATGAGGCGGCGGTACAGAGTGAAGGAAAT 1892
QY 1819 CAGCGTAAGTCAATTTGATCAAGAAAGGAGCGCGGCTCTGTTCAAGACTTCAGGCAATG 1878
DB 1893 GACGCGAGGCTCTACTACATGTGTGATCGGCAATTCGCGTCCCTGAGTTTCAAGCTTGG 1952
QY 1879 AGTGAAGCGCTACGATGTTTTCAACGAGAGGAGGTTCGTAACAGATACCTGACAC 1938
DB 1953 AGCGAGAGCGCACTATGTGTGAACAGAAAGGAGTTGTCACAGAGAACTATACAT 2012
QY 1939 ATCGCAATCAACGAGAGCGCTAAACCTGACAGAGATCTATTAAGCTGTAAAGCT 1998
DB 2013 ATTGCGGTTCACGAGACCGTCTGTGAACAACGAGAGAGAACTACGAAAGATGAGACT 2072

QY 1999 CAGGACACAGACTCAGAAATACGTTCTTGATATTGACGACGAAAGTGTGTTAAGCGAA 2058
DB 2073 GAAAGAACTAGCCGCGAGTACGTTGTCAGTGAATAAAAATGCTGCTCAAGAGAG 2132
QY 2059 GACGAGAGTCCCTTGTGCTTACCGGTGATCTGTGATCTCACCATTTACGAGTTGGC 2118
DB 2133 GAAGCGTGGGGTTGGTGTGGAGAGCTAACCAACCCCGTTCATGAAATGGCC 2192
QY 2119 TACGAGTCTCAGACACGACGACGACCTCAACAAGTCCCAACCATCGAGTCTAT 2178
DB 2193 TACGAGGCTGAAGATCAGGCGCTCGGACCATATTAAGCTACAGTATGAGAGTCTT 2252
QY 2179 GAGTGCAGGTTCAAGTAAATCTGAAATCATCAAAAGCGCTGTGACTTAAGAAAGATCTG 2238
DB 2253 GGGGTTCCGGGATCAGGCAAGTGTGCTATTAATTAAGGCGCTCGTGAACAAACAGATCTG 2312
QY 2239 GTTGTAGTGCAGAAAGAAATCTGGCAGAAATCATCAGGATGTAAAGAGATGAGA 2298
DB 2313 GTCAACGCGGCAAGAGAGAACTGCCAGAAATGTAAACGACGTGAAGAAAGCACCGC 2372
QY 2299 CGTATGATGTGCTGTAGAGCTGTGATTCAGTGTCTTAATGGGGTTAAGCACCCC 2358
DB 2373 GGGAAAGGAGCAAGTAGGAAACAGTACTTCATCTGCTAAGCGGTGCTGTGTC 2432
QY 2359 GTTAACACTCTGTATCATTTGATGAGGCAATTTGCCCTGATCAGGAGACGCTGTGGCACTG 2418
DB 2433 GTGACATGTCTATATGTGAGAGAGGCTTTCGCTTGCAATTCGGGTACTGTGCTGGGCTA 2492
QY 2419 ATTGCAATCTGCAAACT---AGAAAGTGTATTTGTGCGGAGACCCAAACATGCGGC 2475
DB 2493 ATTGCTCTGTATTAACCTCGAGCAAGGTGTGTATGACGAGACCCCAAGCATGCGGA 2552
QY 2476 TTCTTAACTGATGTGCTGAAGATGATCATTTTAACCATGACATATGACATGAGTATC 2535
DB 2553 TTCTTAATATGATGACGCTTTAAGTGAATTTCAACCAACATCTGCACTGAAGTATGT 2612
QY 2536 CATTAAGATCTCTAGAGGTGACACAGACTGTAAACCCGATCGTCTCAAGCTTTC 2595
DB 2613 CATTAAGATATATCACAAGTGTGACGCTGCAAGTGTCAAGCATCTGCACTGAAGTATGT 2672
QY 2596 TACGACAGGAAATGAAGACGTTTAAACCATGTGCTATTAATCATATGATACCA 2655
DB 2673 TACGAGGCAAGATGCCACAGCAACCCGTGCAACCAACCATTAATCAATAGACACCA 2732
QY 2656 GGGACCAAAAGCGGCAACAAAGATGATCTGATTTAACTGTTTCAAGAGTGGTGAAA 2715
DB 2733 GAGACAGCAAGGCCCAAGGAGACATGCTGTTAACATGCTTCCAGGCTGGGCAAG 2792
QY 2716 CAGCTACAGATTTGACTCAAAATACAGAAATCACTGCGGCTGATCGCAAGACTT 2775
DB 2793 CAGCTGCAATTTGACCTACCGTGAACAGAGTATATGACAGAGACATCTCAGGGCTC 2852
QY 2776 ACGCGAAAGCGTTATGTCTGCTAGTACAAAGTCAACAGAAATCACTTACTGCAAG 2835
DB 2853 ACCCGAAAGGGGTATCGCGCTGAAGCAGAAAGTGAATGAATCCCTGTATGACCCCT 2912
QY 2836 ACTTGTAGACGCGAAGGTGTTACTTACAGCAAGAAAGCAATGCTGTGGAAGCG 2895
DB 2913 GCGTGGAGCAAGGAAATGATCTGTGACGCGCACTAGAGATAGCTGTGTGGAAGACG 2972
QY 2896 CTAGCTGTGATCCCTGATTAAGACACTTACAGCTAATATCCCGGGGATTTACAGGCT 2955
DB 2973 CTGGCGGCGATCCCTGATTAAGGTCTATCAAACTTCCACAGGGTAACTTTACGGCC 3032
QY 2956 TCATTTGACGATCTGACGCGGAAACAGACGCACTTATGCGACGCGTTCTTGATTAAGCG 3015
DB 3033 ACATTTGAAGATGCGCAAGAAACAGACAAATATTAAGAGTGAATTAAGAGACCGGCT 3092
QY 3016 CAGACAGTATGTGTTCCAGAAATTAAGTGAAGTGTGCGGGGAGAGGCTTTAGAGCA 3075
DB 3093 GCGCTGTGACGCGTTTCAAGAAACAGCAACGATGTGTGTGGAAGACCTGTGTCT 3152
QY 3076 GTCTTGGCCACGCGCAACATTTGTGCTGACGAGACAGCATGGGAGAC---GTTGACACCA 3132

Db 3153 GTCCGACACTGCGCGAATCAGATTGACAGCAGAGAGTGAGCACCATAATACAGCA 3212
Qy 3133 TTCAGCATGACAGAGCGTACCTACCTGAAATGCGACATGAACTCTTTGACCGAGTTC 3192
Db 3213 TTAAAGAGAGACAGAGCTTACTCTCAGAGTGCGCTTGAATGAAATTTGACCAAGTAC 3272
Qy 3193 TTTGAGTAGACCTGAGACAGTGGGTTAATTTTCGCTCCTAACCGTGCACCTTACTTACAGG 3252
Db 3273 TATGAGTTGACCTGAGACAGTGGCTGTTTTCGCCCCGAGAGTGTCTCTGATTAACAG 3332
Qy 3253 GATGACCATGCGGATTAATCTCGCAGAGGAAACATGATAGGCTTAATAGAGGTAGCA 3312
Db 3333 AACACCATGCGGATTAACAGACCTGTGGAAAGATGATGATTAATGCGGAAACAGCT 3392
Qy 3313 AAGAGTTGTACCGCGCATATCCGTGCATCAAAAAGCGGTGACACAGCGAGGTAGCT 3372
Db 3393 GCCAGGCTGGAAGCTAGACATACCTTCTGAAGGGCGAGTGGCATACGGGCAAGAGCA 3452
Qy 3373 GATATAGGAATTAATACATCAAGACATCTCTCCAAACATTAATGTGGTCCATTAAT 3432
Db 3453 GTTATCGCAGAAAGAAATTCACCGCTTCTGTGCTGAGACATGTAATCTTATCAAC 3512
Qy 3433 CGCCGTTGCCCTCCTGTTGATCGTTGACCAAAAGACAGAGTAACTGATCAAGC 3492
Db 3513 CGCAGGCTGCCGACGCGCTGGTGGTAGTACAAAGACGTTAAAGGAGTAGGGTTGAG 3572
Qy 3493 GATTTCCTATCTAATGAAAGGCAATCTGTGTTGATGATCGCGATCTTATCAGCAT 3552
Db 3573 TGGCTGCTAATTAAGTAAGAGGCTTACCACTGCTGTGTGATGATGATCAACTGGCT 3632
Qy 3553 CCAGGAGAAAGTAGAGTCCATGGGTCCATGGCCCATTAATAC-----CATCAGG 3603
Db 3633 TTGCTCCACGAGAGGTCACTTGTGTTGTACCGCTGATGTACACAGCGCGATAGTGC 3692
Qy 3604 TGTGATCTGATTTGGGAATACCTAGCAGTGTGGTAAATATGACATTAATCTTGTCAAT 3663
Db 3693 TACGACTTAAGTTTAGACTGCGGCTGACGCGGAGGTGTGACTTGGTCTTTGTGAAC 3752
Qy 3664 GTTAGAGACCCGCTACAGAAACATCACTACCAACAGTGCAGAGATCAAGCTATCCACAC 3723
Db 3753 ATTCAACAGGAATTCAGATTCACCACTACAGCAGTGTGTGACACAGCCCATGAAAGTCG 3812
Qy 3724 AGCATGCTAAGCTTAAGGCTGTCCACCACTGAACATGCGGGAACATGTGTGCTATA 3783
Db 3813 CAGATGCTTTGGGGAGATGCGCTACGATCTTAACCGCGGCAATCT--TGTAGA 3869
Qy 3784 GGGTAGGAGCTTGTGATCGCGCAACGAGAAATATCATCATGCGGTGCAAGCTCATTT 3843
Db 3870 GCTTACGAGTACGCGGATTAATTCAGCGAGCGGTTTCTCTTAAGCAAGAAAGTTC 3929
Qy 3844 AGGTTTACCGCTGTCTGTACAGCTTAAGAACTGCGGAAATATGAGAGTCTCTTGGT 3903
Db 3930 TCGTCTGACAGAGTGTGGCGCGGATGTGTACACAGCAATACGAAGTGTCTTGGTGG 3989
Qy 3904 TTCTTGGGCAAGGACCAAGGCAACACACATGACCGAGACAGCTGGGTAGTGGCT 3963
Db 3990 TTCTTCAACTTTGACAAACGAAAGAGACCTTACGCTTACACAGATGATACCAAGCTG 4049
Qy 3964 GACCAATCTTATCAAGGCTCAACAGATACAGAGGAGGAGAGCTCAAGCTACAGAGTG 4023
Db 4050 AGTGCCTGTATGCGGAGAAAGGCAATGCAACGCGCGGTGTGACATCTCAGAGAGTT 4109
Qy 4024 ATCAGAGGTGACATTAAGCAAGAGCGCTGACCAAGCTATGCTTAATGCTTAATGCAAA 4083
Db 4110 AAGAGAGGACATAGCCACGTCACAGAAAGCGCTGTGTGTTAAGCGAGTAAACCCCGT 4169
Qy 4084 GGTCAACAGAGTTCCGAGAGTGTGGGTGACCTGTACCCGAAATGCGCGGCTTTTAT 4143
Db 4170 GGAATCTAGAGGAGTGGCTATGACAGGGCGGTGGCAAGAAATGGCGCTTAAAG 4229
Qy 4144 AGACAGCCAAATGCTGTGCGGACGCGCTAGACTTGTGAAGACGAACCG---CTCATCAT 4200

Db 4230 GGAGCAGCAACACAGTGGGCAATTAATAAACAGTCATGTGCGCTCGTACCCGTCATC 4289
Qy 4201 CATGCTGAGAGACCCAAATTTTCTTAAGTTCGCCGAGCCGAGGGGCACTTAAGTCCGA 4260
Db 4290 CAGCTGAGAGGCTTAATTTCTGTCCACGACTGAAGCGGAAGGGGACCGCAATTTGGCC 4349
Qy 4261 GCTGCTTACATGAGCATAGCTCCATCTGTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db 4350 GCTGTCTACCGGAGAGTGGCCCGCAAGTAAACAGACTGTCACTGAGCAGCTTACCATC 4409
Qy 4331 CCGTACTGTCAACCGCATTTATTTGTGTGCAAAAGATCGAGTATCATTTGAT 4380
Db 4410 CCGCTGTCTCACAGAGTGTTCAGCCGCGAAGAGATAGGCTGCACCAATCTCTCAAC 4469
Qy 4381 CACCTGTCTACTGTTTGTGACATTAAGGATCCGATGTACCATTAATTTGCTGGATAA 4440
Db 4470 CATCTATTCAAGCAATGAGACGCAACGACCTGACGTGACCTTACTGACAGACAA 4529
Qy 4441 CAATGGAGACCAAGATTAATCGAGGCAATTCACCGCAAAAGAAAGCGTCAATTTGAT 4500
Db 4530 AGTTGGAGAAAGAAATTCAGGAAGCATTTGACATGAGAGAGCGCTGTGAGTGTCAAT 4589
Qy 4501 GATGACAAAGCAGTAGATGATGATCTTGTGCAAGGTCACCCAAACAGCTCTTTGCGAGC 4560
Db 4590 GATGACGTGAGCTGACACAGACTTGGTGAAGTGCACCGGACAGCAGCGCTGGTGGT 4649
Qy 4561 AGACAGGTTACTCCGTAAATGAGGCAAGTTGATTAATCTGGAAGTACAGATTC 4620
Db 4650 GCTAAGGCTACAGTACACAGTGGGTGTGATCTGTACTTGAAGTACGAATTC 4709
Qy 4621 CATCACCGCCAAAGACATTCGCAAAATTCATGCAATGTGGCCCAACAAATCTGAGCT 4680
Db 4710 AACAGAGCTGCTATTTGATATGAGAGATTAATGAGATGAGTGTGTGAGTGTGAGAGCA 4769
Qy 4661 AATGACAGATTTGCTGTATCATCTGTGGGAGAGATGTCAAGATCCGCTCCAATGC 4740
Db 4770 AACGACAGATATGCTATATGCGCGCTGGGCAAAACAAATGACATCAGATCCAAATGT 4829
Qy 4741 CCAGTAGAGATCAGAGAGCGTGTCTCCTCCTACACACTTCCATGCTGTGTAATTC 4800
Db 4830 CCGGTAACGATTCGATTCATCAACACTTCCAGAGAGAGTCCGTGTGCGCTAC 4889
Qy 4801 GCTATGAGGCTGAGCGGCTATACAGGTTGGCTCTGCGAAGAAAGACAGTTGCCGTA 4860
Db 4890 GCATGACAGAGAGAGATGCGCGGCTTATGATGATCACCAGTTAAAGCATGTGTT 4949
Qy 4861 TGTCTATCATTTCTGTTCGCAAGATACAGATCAACAGCGTGCAGAGCTACAGTGCAGC 4920
Db 4950 TGTCTATCTTTTCCCTCCGAAATACCATGATGATGGGGTGCAGAAAGTAAAGTGCAG 5009
Qy 4921 AAACAGTCTGTTTTCAGGCGTGTATCACCGCTGTATCACCCAGAGATGACCGGA 4980
Db 5010 AAGTTCTCTGTTTCAGCCGACCGTATCTTCAAGTGTAGTGGCGGAGATATCCCA 5069

RESULT 15
AAK78129
ID AAK78129 standard; DNA; 8010 BP.
XX
XX AAK78129;
XX
XX 20-AUG-1999 (first entry)
XX
XX
DE Semliki Forest virus replicon fragment.
XX
XX Alphavirus; replicon; expression vector; heterologous splice site;
XX RNA splicing; viral RNA replication; antigen; vaccination; therapy; ss.
OS Semliki forest virus.
XX
XX W09925859-A1.
XX
XX 27-MAY-1999.
PD

XX 13-NOV-1998; 98WO-CA01065.
XX 14-NOV-1997; 97US-0065793.
XX (CONN-) CONNAUGHT LAB LTD.
XX Parrington M;
XX WPI; 1999-385140/32.
XX
XX New vector (EV) used to express an antigen for vaccination, e.g.
XX against human immune deficiency virus
XX
XX Disclosure; Fig 3A-F; 60pp; English.
XX
XX This invention describes the construction of a novel expression vector
XX (EV) containing an alphavirus replicon having a heterologous splice site.
XX The expression vector contains (i) a DNA molecule (i) complementary to
XX at least part of an alphavirus RNA genome, and (ii) inserted into a
XX region of (i) that is not essential for its replication, a heterologous
XX DNA (ii), under control of a promoter. (i) includes at least one
XX heterologous splice site (HSS) to prevent aberrant RNA splicing of the
XX alphavirus and is the complement of the complete alphavirus genomic
XX region essential for replication of viral RNA. The expression vector is
XX used to express (ii) in humans or animals, e.g. to express an antigen
XX for vaccination (against human immune deficiency virus) or to produce a
XX therapeutically active protein or peptide. Introducing an HSS: (1) makes
XX it more likely that any splicing will occur at this site, rather than at
XX a cryptic splice site; (2) restores function of the alphavirus when
XX eliminated, and (3) may improve transport of RNA from the nucleus.
XX
SQ Sequence 8010 BP; 2213 A; 2036 C; 2158 G; 1603 T; 0 other;

Query Match 29.5%; Score 1473.8; DB 20; Length 8010;
Best Local Similarity 57.9%; Pred. No. 0;
Matches 2779; Conservative 0; Mismatches 1982; Indels 36; Gaps 8;

Qy 214 ATTTGAAGTGGCGCCGTGACATGACATTCACATCCGCTATCATTCATTCGCTT 273
Db 1 ATCCGACGTGGCCTTCACAGAGAAATGATGTCTACGACAAATACCACTGGTATGCCCT 60
Qy 274 ATGTAAGCGCTGAAGACCCGACAGACTACACGATGACAGAAAGACTTAAAGAAA-- 331
Db 61 ATGGGACGGCAGAAAGACCCGAAAGGCTGATAGCTACGCAAGAAATCGCAGCGGCC 120
Qy 332 ----GTGACTTACCCGACAGAAATGACCTCTTAAAGCGGACAGCTGCTGAAGTCA 387
Db 121 TCCGGGAAGGTGCTGATAGAGATCGCAGAAAAATCACCGACTGACACCGTATG 180
Qy 388 TCAACACACGACGAGAGACTCCATCTCTGTATGACACAGACGCCACGTATAGGTAC 447
Db 181 GCTACGCCAGCGCTGAATCTCTTACCTTTTCTGCTGACACAGAGCTACAGTGTCTAC 240
Qy 448 TTTGGAAGTGAAGATATACCAAGATGTATCCAGTCACTGACCGACATCATCTAC 507
Db 241 GCACCCGAAAGTGGCGGTATACAGAGAGCTGATGTATGATACCAACATCTGCTGAC 300
Qy 508 CACCAAGCGCTTAAAGAGATTAGACAAATTAATGATAGGCTTTGACAGACCCCTTTT 567
Db 301 CATAGGCGATGAAGAGGTGTGAGAAAGCGCTATGATGATGGGTTTGAACCAACCCGTTT 360
Qy 568 ATGTACAAAACATGGGAGGTTCTACCTTACTTACCAACAGAACTGGGCTGAGAGAGA 627
Db 361 ATGTTGACGGCTAGACAGCGCGGTATCAACTACGCAACAAATGGCGAGAGAG 420
Qy 628 GTATTGGAAGACGTAACATTTGGCTCGTAACTCAATCTTACAGAGAGAGGCTTGA 687
Db 421 GTGTTACAGCGCAGAAATAGACTGTGTGACAGATCTTGAATGAGAGAAAGACTGCG 480
Qy 688 AAATCTCAATCTTGAAGAGAGAGGCTCAACCTACTTAATTAAGATCATTTCTGTT 747
Db 481 AAATGTCATTTCTCCGAGAGAAACATTTGAACCTTTCGACACAGTCAATGTTCTCG 540

Qy 748 GGTTCACAAATCTACAGAGATAGATCACTGTTACGTAGCTGGCATCTTCCAAAGTGG 807
Db 541 GGAATCTACATTTGACCTAGAGAGAGAAAGCTACTAGAGAGCTTGACCTTACCTCCGTA 600
Qy 808 TTCCACTTGAAGAGAAAGTCTAACTTACAGAGTATGATGAGAACATTTGCACTGTGAA 867
Db 601 TTCCACTTGAAGAGTAAACAATCTTAACTGATGAGTGCATACATCATATCAATGAA 660
Qy 868 GGGTACCTCATCAAAAGATATACGATACGCCAGAGTATATCGTAAAGTTGAGAACTTG 927
Db 661 GGGTACCTCATCAAAAGATATACGATACGCCAGAGTATATCGTAAAGTTGAGAACTTG 927
Qy 928 GCGTCCAAATGATCGAGAGGTTTCTGATGTTGCAAGTTCACAGATACGCTGCGCGC 987
Db 721 GCGTCCAAATGATCGAGAGGTTTCTGATGTTGCAAGTTCACAGATACGCTGCGCGC 987
Qy 988 GAGAGGGTCTTTTCTGCTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 1047
Db 781 GAAAGAGTCTCATTCCTGTATGACCTAGTCCCTCAACCATCTGTGATCAATGACT 840
Qy 1048 GGGATTTGGCACTGACGTTAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1107
Db 841 GGGATTTGGCACTGACGTTAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1107
Qy 1108 CAAGAGTTGCTCAATGAGAGAGCAAGAAATCTAACAACATGACAGAACTATCTA 1167
Db 901 CAGAGAGTATGTTGAAACGAGAAACAGCAAGCACTAATCAATGAAAGTATCTG 960
Qy 1168 TTACCAAGTGTGCGCCAGGCGTTTTCAGGTGGCGGTGAGTATGTCGCTGCTGAC 1227
Db 961 CTTCCGATGTTGCGCGCTGCGATTTAGCAAGTGGCGGAGGAAATCAAGCAGACCTTGAT 1020
Qy 1228 GAGCAGAAAGAACTAGAGGTTGCGGAGAGCACTTTTATGAGGCTGCTGCTGCTTTC 1287
Db 1021 GATGAAAGAACTTGGGTTGCGGAGAGTATCTTATGAGGCTGCTGCTGCTGCTTTC 1080
Qy 1288 AAGACCCAGAAATCAATCATCTTACAAAGAGCTGTACGCAACAAATTAAGAAAGTA 1347
Db 1081 AAGACCCAGAAATCAATCATCTTACAAAGAGCTGTACGCAACAAATTAAGAAAGTA 1347
Qy 1348 CTTGCGCTCTTGAATCTATTTGATTTCAAGCTTACAGAGCGGCTGATATGAGC 1407
Db 1141 CTTGCGCTCTTGAATCTATTTGATTTCAAGCTTACAGAGCGGCTGATATGAGC 1200
Qy 1408 TTCCGCGTATGAGCTCAAGCTGCTCTTGAACCAACTGTCAAAACCGGCTATATCA 1467
Db 1201 GTCAAGATCAGCAATTAAGATGCTTTTGGCAAGAAAGCAAGCGAG--GTTAATACT 1257
Qy 1468 ATGACCGATGTGAGCACTGTGCTGCTTACAGCAAGAACTGAAAGTGTGACGCG 1527
Db 1258 GTTCTGACCGCTGTGATGAGCAGAGAGTGTGAAACAGAGAGAGAGAGTGTGAGGCC 1317
Qy 1528 GAAGAGATCAGAGAAAGCTTGCACCTCTGCTCTGAAATGAGAAAGAGACCTGAG 1587
Db 1318 GAGCTGACTAGAGAGCTTATACCACTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1377
Qy 1588 GCAGAAATTAACCTCATTTATGCAAGAG-----GCAGAGCAGGTACGCTGAGAGCA 1641
Db 1378 GTGACGCTGACGTTGAAGATCTAGATATCAAGCAGAGTGAAGAGTGTGAGAAACCT 1437
Qy 1642 GCAGAGACATCAAGGTTGCAAGTTACCCAGGCGAAGAGAAATGAGGTCTTACGCTATA 1701
Db 1438 GCGAGCGCTGTTAAAGTACACGACAGCCAGACAGAGTACTATGAGAAATTAAGTATG 1497
Qy 1702 CTTTCACTCCAGGCGGTATTAATGATGAAATTAAGTAACTGCTATGACCGGAGAA 1761
Db 1498 CTGTCCCGAGACCGTGTCTCAAGAGCTCAAGTGTGCGCCGCTGCTGAGAGAG 1557
Qy 1762 CAAGTACTGTTATGATCTCAAGAGTGTGAGAGAGTATCAAGAGTGTGAGAGTATCAC 1821
Db 1558 CAGGTGAAATTAATTAACATTAACAGGAGGCGGCGTTTACAGAGTGTGAGAGTATGAC 1617

1832 GGTAAAGTCAATTGTAACGAAAGGACGGCGGTCCTGTTCAAGACTTCCAGGCATGTAGT 1881
1618 GGCAAGGAGTCTTACTACTATGTGGATTCGGGCTTCCTGAGTTTCAAGCTTTGAGC 1677
1882 GAGACCGCTACGATCGTTTTCACAGAGAGGAGTTCTGTAACAGATCTCGACCAATC 1941
1678 GAGACCGCTACGATCGTTTTCACAGAGAGGAGTTCTGTAACAGATCTCGACCAATC 1737
1942 GCAATCAACGAGAGCGCTTAACACTGACGAAAGTACTATTAAGACTGTAAAGACTCG 2001
1738 GCCGTTACGACCGCTCGCTGTAACACGAGAGAGAACTACGAGAAATCAGAGACTGA 1797
2002 GACACAGACTCAGAAATACGCTCTTCAATTTGACGACGAAAGTGTGTTAAGCGAGAA 2061
1798 AGAATGACGCGCGATACGTTGTTTCAACGATTAATAAAATGCTGCGTCAAGAGAGGA 1857
2062 GCAGGTCCCTTGTGCTTAACCGGTATCTGTAAGATCCACCTTTCAGAGTTTGGTAC 2121
1858 GCGTCGGGTTGTGTGTGGAGAGCTAACCAACCCCGCTTCATGAAATTCGCTAC 1917
2122 GAGAGTCTCAAGACACGACCAAGCAGACCTCACAAGTCCCAACCATCGAGTCTATGGA 2181
1918 GAAAGGCTGAAAGATAGGCGGCTGCGCACCATTTAAGATCACTAGTGTAGAGTCTTTGG 1977
2182 GTGCCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGCTGTGACTTAAGAAAGATCTGGT 2241
1978 GTTCCGGGATGAGCAAGTCTGCTATTTAAGAGCTGTGTTCCAAACAGATCTGGTC 2037
2242 GTGAGTGGAGAGAGAAATCTGCGCAAAATCATCAGGAGTGTAAAGAGATGAGAGCT 2301
2038 ACCACCGGCAAGAGAGAACTGCCAGAAATATGTTAAAGAGTAAAGAGAGACCGGG 2097
2302 ATGAGTGTGCTGCTAGAGCTGTGATTCAGTCTTCAATTAAGGGGTTAAGACCCGCT 2361
2098 AAGGGCAAGATGAGGAAACAGTACCTCATCTGCTTAACGAGGTTGCTGTGCGG 2157
2362 AACACTGTATCATTTAGTAGAGGCAATTTGCTGCTCAATGACGAGCGCTGTGCACTGAT 2421
2158 GACATCTTATATGTGAGAGAGGCTTTCGCTTCCATTCGGTATCTGCTGCGCTTAAT 2217
2422 GCCATCTCAAACT--AAGAAAGTGTATTTGCGGGGAGCCCAAAATATGCGGCTTC 2478
2218 GCTCTTGTAAACCTCGGAGCAAGTGTGTATCGGAGACCCCAAGCAATGCGGATTC 2277
2479 TTTAATCATGATGCTGCTGAAAGTACATTTTACCAATGACATATGAGTGAATACAT 2538
2278 TTTCAATATGATGAGCTTAAGGTGAATTTCAACCAACATCTGATGAGTATGAT 2337
2539 AAAAGCATCTTAGAGAGTGCACAGACTGTAAACCGGCATGCTCCACGCTCTTCTAC 2598
2338 AAAAGTATATCAGAGCTTTCACGCGTCACTGCAAGGCAATGTGTCTTACGTTGACATC 2397
2599 GACAAAGCAATGAGAGCGTTTAACTGCTGTCTGATTAATAATCATATAGATACACAGGG 2658
2398 GGAGGCAAGATGCGACGACCAACCGCTGCAAAACCATATATATAGACACACAGGA 2457
2659 ACCAAGAGCGGCAAGAGATGATCTGATTTCTAATCTGTTTCAAGAGATGGGTGAACAG 2718
2458 CAGACCAAGCCCAAGCAGAGACATCGTGTAAATCTTCGAGGCTGCAAAAGAG 2517
2719 CTACAGATTTGATCAAAATATCAGAAATCATCTGCGGCTGATCCGCAAGAGCTTACG 2778
2518 CTGAGTTTGAATCCGTGACACGAGATCATGACGAGACGATCTCAGGAGCTTCAAC 2577
2779 CGGAAAGCGTTTATGCTGTCAAGTAAAGTCAACGAGATCACTCTACTGCGAGACT 2838
2578 CGCAAGAGGATATAGCGCGTAAAGGCAAGGTGAATAATCCCTTGATGCCCTGCG 2637
2839 TCTGAGCAGTGAAGTGTATCTTACAGGCAAGAAACGATGTCTGGAAGAGCGTAA 2898
2638 TCGGAGCAGTGAAGTGTATCTGACGCGCACTGAGATAGGCTGTGTGGAAGAGCGTGG 2697
2899 GCTGTGATCCCTGATAAAGACACTTACAGCTAAATATCCGCGGATTTTCAGCGCTTCA 2958

2698 GCCGCGATCCCTGATTAAGGTCTTATCAAACTTCCACAGGAGTAACTTTACGCGCACA 2757
2959 TTGACGACTGCGACCGGGAACAGACCCATTAATGACGCGTTCTTGAATAGCCGAG 3018
2758 TTGAGAGATGCAAGAGAAACAGCAAAATTAAGAGGTGATTAAGAGCCGCTGCG 2817
3019 ACAGCTGATGTGTTCCAGATTAAGGTGAACGCTGCTGGCGGAAGCTTTAGAGCAGTC 3078
2818 CCGTGTGACGCGTTCCAGAAACAAAGCGAAGCTGTGTTGGGGAAGAGCTGTGCTGTC 2877
3079 TTGGCAGCGCCAACTTGTGCTGACGAGACAGAGTGGAGAC--GTGCAACCATTC 3135
2878 CTGACACTGCGCGGAATCAGATTGACACAGAGAGTGGACACCATTAATTAACACTTT 2937
3136 AAGCATGACAGAGCTACTACCTGAAATGACCACTGAACCTTTTGTGACCAAGTCTTT 3195
2938 AAGAGGACAGAGCTTACTCTTCAGTGTGCGCTTGAATGAATTTGACCAAGTACTAT 2997
3196 GAGATGACCTGACAGTGGGTTATTTCCGCTCCTACCGTGCACCTTAAGTACAGGAT 3255
2998 GAGTTGACCTGACAGTGGCTGTTTCTGCTCCGAAAGTGTCTGTATTTACAGAAC 3057
3256 CAGACTGGGATTAATCTGCCAGGGAAGACATGATGGGCTTAATAGAGGTAGCAAG 3315
3058 AACCACTGGATTAACAGACTGTGGAGAGATGATGATTTCAATGCCGAACAGCTGCC 3117
3316 GAGTTGTACGCGGATATCCGTGATCACAAAAGCGTTTGACACAGGAGGTACTGAT 3375
3118 AGGCTGAAGGTAGACATATCTTCTGAAAGGCGAGTGGCATACGGGCAAGCGAGTT 3177
3376 ATTAAGAAATTAATPACATCAAGAGATCTCTCAACAAATTAATGCTTCCATTAATGC 3435
3178 ATCCGAAAGAAATAATCAACCGTTTCTGTGCTGCAATATGATTTCTATCAACGC 3237
3436 CGGTTGCCCACTGTTGATGCTTGAACCAAGAGACAGGTACAACTGATCACAGCGGA 3495
3238 AGGCTGCCGACCGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3297
3496 TTCTATATTAAGAGAGGCAATCTGTGTGTGATGATGATGATGATGATGATGATGAT 3555
3298 CTGCTCATTAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3357
3556 GGAAGAAAGTAAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3606
3358 CCGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3417
3607 GATCTGATTTGGGATTAATCTAGGCAATGCTGATTAATTAATGATTAATTTGTCATGTT 3666
3418 GACCTAAGTTTAAAGATGCTGCGGCTGACGCGGAGGTTCGACTGTGCTTTGTGAACAT 3477
3667 AGGACCCGCTACAGAACTATCACTACAGATGATGATGATGATGATGATGATGATGATG 3726
3478 CACACGAAATTAAGATTAACACTACAGAGATGATGATGATGATGATGATGATGATGAT 3537
3727 ATGCTAAGTAAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3786
3538 ATGCTGAGGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3594
3787 TATGCTGTTGCTGATTCGCGCAACCGAATATCATCTGCGGTGCGACGCTCATTTAGG 3846
3595 TACGATATGCGGATTAATTAATGAGGAGCGCTTGTCTCTTTAAGCAAGAAATTCG 3654
3847 TTTACCGCTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3906
3655 TCTGCAAGAGTGTGCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 3714
3907 TTGCGCAAGGACAGAGCAACCAACACATGACAGGACAGACTCGGTGATGATGATGATGAT 3966
3715 TCAACTTTGACAGGAGAGAGACCTCTACGCTACCAAGATGATTAACAGGTGATGAT 3774
3967 AACATTAATCAAGGATCAACAGGTACAGAGCAGGAGAGCTCCAGCGTACAGATGATC 4026

Db 3775 GCCGTGTATGCGGAGAACCATGACACGCGCGGGTGTGACCATCTTACAGATTAAAG 3834
QY 4027 AGAGGTACATTACAGAGAGCGCTGACCAAGCTATCGTTATGCTTAATAGCAAGGT 4086
Db 3835 AGAGCAGACATAGCCACGTGACAGAAAGCGGCTGTGTAAAGCAGCTAACGCCGTGGA 3894
QY 4087 CAACCAAGTTCCGGAGTGTGCGGTGACACTGACGAAAAATGGCCGGCTGCTTTTGAATGA 4146
Db 3895 ACTGTAGGGGATGCGGTATGAGAGGCCCGTGGCGAAAGAAATGGCCGTCAAGCTTTAAGGA 3954
QY 4147 CAGCCAAATAGCTGTGGGACGGCTAGACTGTGAAGCAGAACCG--CTCATCATACAT 4203
Db 3955 GCAGCAACACCAAGTGGGACCAATTAAAAAGTCAATGTGCGGCTGTACCCCGTCATCCAC 4014
QY 4204 GCTGTAGAGACCCCAATTTTCTTAAGTGTCCGGAACCGAGGGCGACCTTAAGCTGGCAGCT 4253
Db 4015 GCTGTAGCGCTTAATTTCTTGCCACGACTGAAGCGGAAGGGGACCGGAAATGGCGCT 4074
QY 4264 GCCTACATGAGCATAGCGTCATCGTCACGCTGAGCGGATTACAAAAATATCAGTACCG 4323
Db 4075 GTCTACCGGGGAGTGGCCCGCGGAAGTAACAGACTGTCTACAGCAGGTAGCCATCCG 4134
QY 4324 CTACTGTCAACCGGCATCTATTTGTGTGCAAGATGAGTGAATCAATTCATTCAC 4383
Db 4135 CTGCTGTCCAGAGAGTGTGAGCGGCGAAGAGATAGGCTGACGAATCCCTCAACCAT 4194
QY 4384 CTGTTCACTGCTTTCGACACTAGCGGATGCCGATGTCACCATATATGCTTGGATTAACAA 4443
Db 4195 CTATTCAAGAAATGAGGACGACGACGATGACGATCTACTGACAGACAAAGT 4254
QY 4444 TGGAGACCAAGATATGAGGCAATTCACCGCAAAAGCAATTCGAAATTCGAGATGAT 4503
Db 4255 TGGGAGAAAGAAATTCAGGAAGCCATTGACATGAGAGCGGTGTGAGTTGCTCATGAT 4314
QY 4504 GACAAGCCAGTAGAATTGACTTGTGAGGCTCCACCACCAACAGCTCTTTGGCAGGCGA 4563
Db 4315 GACGTGAGCTGACACAGATGTTGTGAGTGCACCGGACAGCAGCTGTGTGCTGT 4374
QY 4564 CCAGGTTACTCCGTCATGAGGGAAGTGTATTCATCTGGAAGGTAACGATTCAT 4623
Db 4375 AAGGCTACAGTACACTGACGAGGTCGCTGATCTGTACTTGAAGGTACGAAATTCAC 4434
QY 4624 CAGACCCCAAGACATTTGCCGAATCCATGCAATGTGCCCCCAAAATCTGAGGCTAAT 4683
Db 4435 CAGGCTGCTATGATATGCGAGATACTGACGTTGTGCCCAAGCTGCAAGAGCAAAC 4494
QY 4684 GAGCAGATTGCTGTACATCTTGGGAGAGATGTCAGCATCCGCTCCAAATGCCCA 4743
Db 4495 GAAACAGATATGCTATAGCGCTGGGCGAAACAATGACACATCAGATCCAAATGTCCG 4554
QY 4744 GTAGAGAGTGAAGGCGTCTGCTCCACTCACACTTCCATGCTGTGTATTAATTCGCT 4803
Db 4555 GTGAACGATTCGATTCATCAACCTCCAGGACAGTGCCCTGTGTGCCGCTACGCA 4614
QY 4804 ATGACGCTGAGCGCGTATACAGGTTGGCTCTCGAAGAAAGAACAGTTCCCGTATGC 4863
Db 4615 ATGACAGCAAGACGATCGCCCGCTTAGTCAACACAGTTAAAGCATGTGTGCTTGC 4674
QY 4864 TCATCATTTCTGTTGCCGAATGACAGATCAAGGCGTGCAGAAAGCTACAGTGCAGCAA 4923
Db 4675 TCATCTTTTCCCTCCGAAATACATGTAGTGGGGTGCAGAAAGTAAAGTGCAGAGAG 4734
QY 4924 CCACTCTGTTTTCAGGCGTGTACCAACCGGCTGTACACCCCAAGAAATACCGCGAA 4980
Db 4735 GTTCTCTGTTGACCCGACGATCTTCAGTGTGTAGTCCGCGAAGTATGCGCA 4791

Sindbis virus cDNA
alphavirus-based euk
Sindbis-like viruses
Representative euk
Representative euk
Representative euk
Representative euk
Representative euk
Eukaryotic layered
Eukaryotic layered
Sindbis-like viruses
DNA of expression
Sinchitronp virus
Sindichiton virus
Polynucleotide seq
Plasmid pTES/2J DNA
Venezuelan equine
Venezuelan equine
Polynucleotide seq
Plasmid BGD-EB DNA
Polynucleotide seq
987Bneo RNA sequence
PSP6-SFV4 RNA tran
Venezuelan equine
Human immunodeficit
Venezuelan equine
Venezuelan equine
Venezuelan equine
Sindbis virus cDNA
Sindbis virus vari
Sindbis virus vari
Sindbis virus cDNA
Polynucleotide seq
Sindbis expression
Alphaviral vector
Nucleotide sequenc

PT encephalitis virus genome - useful for the production of live or
PT attenuated vaccines for human or veterinary medicine

PS Example 1; Page 54-67; 112pp; English.

XX This CDNA sequence codes for an infectious western equine
CC encephalitis (WEE) virus RNA transcript. DNA representing the
CC entire genome was prepared by PCR using primers (see AA74110-21)
CC based on partial genome sequences. 5' Sequences were obtained by
CC RACE. The full-length infectious clone is useful in the production
CC of virulent WEE virus, and for introducing and testing attenuating
CC mutations. Also new are: infectious or attenuated WEE RNA
CC transcripts and WEE viral particles; cDNA (see AA74108) encoding an
CC infectious Venezuelan equine encephalitis (VEE) variant IE viral
CC genome; infectious or attenuated VEE RNA transcripts and viral
CC particles; an attenuated chimeric virus containing non-structural
CC sequences from a first alpha-virus (AV) and structural sequences
CC from a second AV, resulting in attenuation of the second AV; a method
CC for expressing a protein (especially an antigen to protect against
CC a pathogen) by cloning its gene into an attenuated WEE or VEE IE
CC replicon so that transcription of the replicon produces RNA able to
CC infect the cells in which protein is to be produced; methods for
CC diagnosing WEE and VEE infections; and polypeptides encoded by VEE
CC variant IIA. Attenuated WEE and VEE are used in live or inactivated
CC vaccines, for use in human or veterinary medicine. Chimeric viruses
CC are also useful as vaccines, directed against the second AV which is
CC particularly Eastern equine encephalitis (EEE) virus or some variant
CC of VEE. WEE and VEE nucleic acids are used as primers and probes to
CC diagnose virus infections and to define natural variants, also for
CC production of protein antigens which can be used as diagnostic
CC reagents, to generate antibodies, and in vaccines. The attenuated
CC viruses are highly immunogenic and provide long-lasting protection.

XX Sequence 11492 BP; 3257 A; 2854 C; 2794 G; 2570 T; 17 other:

Query Match 90.6%; Score 5877.6; DB 20; Length 11492;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 6238; Conservative 0; Mismatches 221; Indels 29; Gaps 20;

QY 1 CACCGCCAGCAGCAACCGTAATATGTGAACCACTGTGCCAGAACGTATACCCAGTC 60
DB 5012 CATGCCAACAACCAACCTTAATATGTGAACCAACTGTGCCAGAACGTATACCCAGTC 5071
QY CGGTGATTTCTAGAGCAACAGTGGCGAATGACTGCTATGCTGGGGCGCTCGTCTT 120
DB 5072 CGGTGATTTCTAGAGCAACAGTGGCGAATGACTGCTATGCTGGGGCGCTCGTCTT 5131
QY 121 CTAGCTGTGCACACGCTGCTCAACCGCTGGAGCGACTATAGACAGCGGTTGTGTTA 180
DB 5132 CTAGCTGTGCACACGCTGCTCAACCGCTGGAGCGACTATAGACAGCGGTTGTGTTA 5191
QY 181 CAGCTGATGTGCATCAAGCGAACGCTCTACGTGAGAGATCCCTAGTGTCTTGGCTTGG 240
DB 5192 CAGCTGATGTGCATCAAGCGAACGCTCTACGTGAGAGATCCCTAGTGTCTTGGCTTGG 5251
QY 241 ACGTCCAGCTGCTTGAAGCTCACTGATTCCTGAGATTCCTGAGTGCACAGAGCT 300
DB 5252 ACGTCCAGCTGCTTGAAGCTCACTGATTCCTGAGATTCCTGAGTGCACAGAGCT 5311
QY 301 TTGAAGTGAAGAACCACTCTGTACAGACCTTAACGTGCGAGATGCGAAGCTGTGGAC 360
DB 5312 TTGAAGTGAAGAACCACTCTGTACAGACCTTAACGTGCGAGATGCGAAGCTGTGGAC 5371
QY 361 TGGCGCAATATATCAAGACTTCAATATCTGCTCTTTCAGATTTCTTTGGACTACAGAC 420
DB 5372 TGGCGCAATATATCAAGACTTCAATATCTGCTCTTTCAGATTTCTTTGGACTACAGAC 5431
QY 421 CAGTACCGGACCAACGAGAGAGCCCATCCATCACTTGAATGCAAGGCTTCCGCACTTC 480
DB 5432 CAGTACCGGACCAACGAGAGAGCCCATCCATCACTTGAATGCAAGGCTTCCGCACTTC 5491
QY 481 CAGTCCAAAGCCAGCAGACTAAGTACCAACAACCAAGAGTGCCTGAGACGATCT 540

DB 5492 CAGTCCAAAGCCAGCAGAACTAAGTACCAACAACCAAGAGTGCCTGAGGCGATCT 5551
QY 541 CAGAACGGAGTTGAGAGAGTACATCGTCAACACTTCAACTGACGGTATGAAGCGGAG 600
DB 5552 CAGAACGGAGTTGAGAGAGTACATCGTCAACACTTCAACTGACGGTATGAAGCGGAG 5610
QY 601 CGTATATTTCTCATCGAAGAACAGGCCCAAGTCACTTCAAGAGAAATCAGTACGTCAAT 660
DB 5611 CGTATATTTCTCATCGAAGAACAGGCCCAAGTCACTTCAAGAGAAATCAGTACGTCAAT 5667
QY 661 GTAAATACAAAGAACCTATATGTGATCGGCGCTCCATGAGAAATTAAGCCCGCGCC 720
DB 5668 GTAAATACAAAGAACCTATATGTGATCGGCGCTCCATGAGAAATTAAGCCCGCGCC 5726
QY 721 TCGATCTCGAAGAGAGAAATGTTACAGAAAGAACTGCAATTAATGCGCTTC-TGAAGGA 779
DB 5727 TCGATCTCGAAGAGAGAAATGTTACAGAAAGAACTGCAATTAATGCGCTTCGTAAGGA 5786
QY 780 AATGAAAGCAGTATCAATACAGAAAGTA-GAAATATGAAAGCAATTAACGCGAGCG 838
DB 5787 AATGAAAGCAGTATCAATACAGAAAGTAAAGTAATTAAGAAATTAACGCGAGCG 5846
QY 839 ACTCATTTCTGATTTGGGCAATATATCATATCATGAGAAATCTGTGAGTGTACAG 898
DB 5847 ACTCATTTCTGATTTGGGCAATATATCATATCATGAGAAATCTGTGAGTGTACAG 5906
QY 899 AGTCAATATTCCTGTACCAATCTACTGTCAACGGTATTAACAGTTTACATCTGACA 958
DB 5907 AGTCAATATTCCTGTACCAATCTACTGTCAACGGTATTAACAGTTTACATCTGACA 5966
QY 959 GGTGCGGGTAAACGTCGCACTTAAGTATCAAGAGAAATTAACCTTCACTAGGCACTTA 1018
DB 5967 AGTGGCGGTAAAGTTTGCACTTATGATCTCAAGAGAAATTAACCTTCACTAGGCACTTA 6026
QY 1019 TTGTATTAACAGATGAATACGATCGATCTGTAATGATGAGAGCGGCGCATCTGTCTCT 1078
DB 6027 TTGTATTAACAGATGAATACGATCGATCTGTAATGATGAGAGCGGCGCATCTGTCTCT 6084
QY 1079 AGATACAGCCACTTTTGTGCGGCTAACTGAGAAAGTACCAGAAAGAGATGCTATT 1138
DB 6085 AGATACAGCCACTTTTGTGCGGCTAACTGAGAAAGTACCAGAAAGAGATGCTATT 6140
QY 1139 GCAGCCAGAGATGAATACGATCGATCTGTAATGATGAGAGCGGCGCATCTGTCTCT 1198
DB 6141 GCAGCCAGAGATGAATACGATCGATCTGTAATGATGAGAGCGGCGCATCTGTCTCT 6200
QY 1199 GGTGCGGCTAAAGAAATGCAACGTTACCCAAATGCGAGAAATTAACCTGTCTTGA 1258
DB 6201 -GCTGACGCACTTAAAGAAATGCAACGTTACCCAAATGCGAGAAATTAACCTGTCTTGA 6257
QY 1259 TTGGCGGCGCAATTAATGATGATGTTTCAAGAAATACGCAATGCAATGAGTACGTGG 1318
DB 6258 TTGGCGGCGCAATTAATGATGATGTTTCAAGAAATACGCAATGAGTACGTGG 6317
QY 1319 TACCTTTCGCGATTAACCTTATGCGCTTAACCTAAGAGAAATGAGTATGAGCAAA 1378
DB 6318 TACCTTTCGCGATTAACCTTATGCGCTTAACCTAAGAGAAATGAGTATGAGCAAA 6377
QY 1379 GCTGAAAGGCGCGAAGCAGACATTTTGGCAATATCTCATATCTTAAACCGTTGCA 1438
DB 6378 GCTGAAAGGCGCGAAGCAGACATTTTGGCAATATCTCATATCTTAAACCGTTGCA 6437
QY 1439 GAGATATCAATGATTAATTTGCTCATGATCTTAAAGAGATGTCAAAGTTACTCCCGG 1498
DB 6438 GAGATATCAATGATTAATTTGCTCATGATCTTAAAGAGATGTCAAAGTTACTCCCGG 6496
QY 1499 CAGGAAACATACAGAGAGCGGCTAAGTGCAGGTTATTCAGAGTGCAGATCCCTTGC 1558
DB 6497 CAGGAAACATACAGAGAGCGGCTAAGTGCAGGTTATTCAGAGTGCAGATCCCTTGC 6556
QY 1559 TACCGCTTACCTTTGCGGAGTCCATCGGGAATTAATGCTGTAAGTCTGATGCTTCT 1618
DB 6557 TACCGCTTACCTTTGCGGAGTCCATCGAGAAAT--GTCCGTAAGTCTGATGCTTCT 6614

QY 1619 GCCAAATATCCATACCTCTCTTCGACATGTGACGGAGATTGATGCGATTATGCTGA 1678
Db 6615 GCCGAATATCCATACCTCTCTTCGACATGTGACGGAGATTGATGCGATTATGCTGA 6674
QY 1679 ACATTTCCACACGGCGACCCAGTATGGAACCGACATCGCGTCTTTGATTAAGCGA 1738
Db 6675 ACATTTCCACACGGCGACCCAGTATGGAACCGACATCGCGTCTTTGATTAAGCGA 6734
QY 1739 AGACGAGCTATCGCATTTGCGGGTGTGATGATCCTTAGAGACTTAGGTGCGACCAAC 1798
Db 6735 AGACGAGCTATCGCATTTGCGGGTGTGATGATCCTTAGAGACTTAGGTGCGACCAAC 6794
QY 1799 GCTTTAGATTGATAGAGCGCGCTTCGCGCAATATCAATCTGTGCACTTACCTACAG 1858
Db 6795 GCTTTAGATTGATAGAGCGCGCTTCGCGCAATATCAATCTGTGCACTTACCTACAG 6854
QY 1859 AACGAGTTTAAATTTGGTCCATGATGAATCCGGTATGTTCTTAAACGCTGTTTCA 1918
Db 6855 AACGAGTTTAAATTTGGTCCATGATGAATCCGGTATGTTCTTAAACGCTGTTTCA 6914
QY 1919 CACACTAGTCAATATCATGATTTGCTAGAGAGTACTAGTGAACGGTTAAACAGCTGAC 1978
Db 6915 CACACTAGTCAATATCATGATTTGCTAGAGAGTACTAGTGAACGGTTAAACAGCTGAC 6974
QY 1979 GTGCGCGGCTCTATCGGCGACGATATAGTGTGATGCTGCTCCGACACTTGTAT 2038
Db 6975 GTGCGCGGCTCTATCGGCGACGATATAGTGTGATGCTGCTCCGACACTTGTAT 7032
QY 2039 GCGCGAGATGCGGCACTTGGCTGACATGAAATGAAATTAATTTGATGCACTTATGG 2098
Db 7033 GCGCGAGATGCGGCACTTGGCTGACATGAAATGAAATTAATTTGATGCACTTATGG 7092
QY 2099 TATCAAAACACCTTATCTGTGGGGATTTATCTGTGTGACGAGTAAACGACGACG 2158
Db 7093 TATCAAAACACCTTATCTGTGGGGATTTATCTGTGTGACGAGTAAACGACGACG 7150
QY 2159 CTGCGAGTCCGACGCTCTTAAAGAGCTTTTAAAGCTTGAAGAAACCATTCGAGTGA 2218
Db 7151 CTGCGAGTCCGACGCTCTTAAAGAGCTTTTAAAGCTTGAAGAAACCATTCGAGTGA 7209
QY 2219 TGATACCCAGACCTGCGACCGCGCGGCACTGCTGATGATGAAGCAATGCGATGGAAC 2278
Db 7210 TGATACCCAGACCTGCGACCGCGCGGCACTGCTGATGATGAAGCAATGCGATGGAAC 7269
QY 2279 AATTGGAATTAACGAGAGTATAGTAAGGCGGTAAATCCAGATATCAATCTGCG 2338
Db 7270 AATTGGAATTAACGAGAGTATAGTAAGGCGGTAAATCCAGATATCAATCTGCG 7329
QY 2339 AGGCTGATCATCACTCTGTCCACGTTAGCGGAAAGCGTTAAGAACTTCAAGAGAT 2398
Db 7330 AGGCTGATCATCACTCTGTCCACGTTAGCGGAAAGCGTTAAGAACTTCAAGAGAT 7389
QY 2399 AAGAGGAGCCCAATCACCTCTTACGCTGACCTTAAATAGGTGACGTAGTGAACGAC 2458
Db 7390 AAGAGGAGCCCAATCACCTCTTACGCTGACCTTAAATAGGTGACGTAGTGAACGAC 7449
QY 2459 CTACCCACCGCG-AGAATGTTTCCATACCTTCAAGCTGAACCTTCCACAGTTTACCTAC 2517
Db 7450 CTACCCACCGCGCAAAATGTTTCCATACCTTCAAGCTGAACCTTCCACAGTTTACCTAC 7509
QY 2518 AAATCCGATGGCTTACGAGATCCAAACCTCTTGGGGCGGCTGGAAGCGTTTGGGGC 2577
Db 7510 AAATCCGATGGCTTACGAGATCCAAACCTCTTGGGGCGGCTGGAAGCGTTTGGGGC 7569
QY 2578 CCCGCTGCGCTCAATCGAAGATCTTAGAGAGTGAATGATCACTTGAATTTCAACA 2637
Db 7570 CCCGCTGCGCTCAATCGAAGATCTTAGAGAGTGAATGATCACTTGAATTTCAACA 7629
QY 2638 ACGATCACTTAATCCGCGCGGAGTCCACCGCAAGAAAGAAAGAGTGTCTTAAAGC 2697
Db 7630 ACGATCACTTAATCCGCGCGGAGTCCACCGCGCAAGAAAGAAAGAGTGTCTTAAAGC 7689

QY 2698 AAAACCTATCAGCTTAAAAAGAAAGACAGCAAGCCAGAGAGACGAAACCAAGCTTAA 2757
Db 7690 AAAACCTATCAGCTTAAAAAGAAAGACAGCAAGCCAGAGAGAGACGAAACCAAGCTTAA 7749
QY 2758 ACCAGGAAACGACAGATATGTGTATGTAAGTGTGGAAGTCCGACAAAGACATTTCCGATCAT 2817
Db 7750 ACCAGGAAACGACAGATATGTGTATGTAAGTGTGGAAGTCCGACAAAGACATTTCCGATCAT 7809
QY 2818 GCTGAACGCGCAAGTGAATGATATGCTGTGCTGCGAGAAAGGCTGATGAACCACT 2877
Db 7810 GCTGAACGCGCAAGTGAATGATATGCTGTGCTGCGAGAAAGGCTGATGAACCACT 7869
QY 2878 CCACGTTGAAGAAAAATTTGATTAATGACATTAAGCGGCGCTGAAATTTGAAGAGGCTTAG 2937
Db 7870 CCACGTTGAAGAAAAATTTGATTAATGACATTAAGCGGCGCTGAAATTTGAAGAGGCTTAG 7929
QY 2938 CATGTAGCATTTGGAATGAGCGGACGTTCCCGAAGACATGAATTCAGACGCTGCAAGTA 2997
Db 7930 CATGTAGCATTTGGAATGAGCGGACGTTCCCGAAGATGAATTCAGACGCTGCAAGTA 7989
QY 2998 CACGAGGAGCAACACCGGCTCTCAACTGGGACACAGGGCGAGTCCAGTATGAAGAA 3057
Db 7990 CACGAGGAGCAACACCGGCTCTCAACTGGGACACAGGGCGAGTCCAGTATGAAGAA 8049
QY 3058 TGGAGATTTTACCGTACCGAGAGAGTGGGCGGAAAGGCGACAGCGGAAACCGATCCT 3117
Db 8050 TGGAGATTTTACCGTACCGAGAGAGTGGGCGGAAAGGCGACAGTGGAAACCGATCCT 8109
QY 3118 GGAACAACAGGACAGAGTGTGCTATTTGTTCTAGAGGTGCAATGAGGCAACGCTTAC 3177
Db 8110 GGAACAACAGGACAGAGTGTGCTATTTGTTCTAGAGGTGCAATGAGGCAACGCTTAC 8169
QY 3178 GGGCCTTTCAAGTGTCACTTGGAAACGAAAGGGGTGACATCAAGATATCCCCGAAG 3237
Db 8170 GGGCCTTTCAAGTGTCACTTGGAAACGAAAGGGGTGACATCAAGATATCCCCGAAG 8229
QY 3238 TTCTGAACCGTGTCACTAGTATGACGCGTATGCGGCTTTCGATGTCAAGTCCCATG 3297
Db 8230 TTCTGAACCGTGTCACTAGTATGACGCGTATGCGGCTTTCGATGTCAAGTCCCATG 8289
QY 3298 CGAACAACACCGTGTGCTATTTCACTGACCGCAAGAAACACTCGAGTGTGGAAGA 3357
Db 8290 CGAACAACACCGTGTGCTATTTCACTGACCGCAAGAAACACTCGAGTGTGGAAGA 8349
QY 3358 GAAAGTGAACATCAATTAACGACGCTGTGGAAGACGCTTGAATTCATCAACG 3417
Db 8350 GAAAGTGAACATCAATTAACGACGCTGTGGAAGACGCTTGAATTCATCAACG 8409
QY 3418 CCGGCCCCAAGAGCAATTAACGATGACTTCACTGACCGATCCCTTCTGGGTTCTG 3477
Db 8410 CCGGCCCCAAGAGCAATTAACGATGACTTCACTGACCGATCCCTTCTGGGTTCTG 8469
QY 3478 CCGGCTATTCAGACACTCAAGCGCGTGTTCAGGCCCAATAAAAATGGAACGTTGGGA 3537
Db 8470 CCGGCTATTCAGACACTCAAGCGCGATGTTTAGCCCAATAAAAATGGAACGTTGGGA 8529
QY 3538 CGAATTCGATGATGATGATGATTAAGATTCAGAGTCTCGGCACAATTCGCTCAATCAGGC 3597
Db 8530 CGAATTCGATGATGATGATGATTAAGATTCAGAGTCTCGGCACAATTCGCTCAATCAGGC 8589
QY 3598 AGGCACTGCGGATGCAACAAATTCGCTTACATGCTTTGACACAGCAACGATGACATGAA 3657
Db 8590 AGGCACTGCGGATGCAACAAATTCGCTTACATGCTTTGACACAGCAACGATGACATGAA 8649
QY 3658 GGAAGAAGTATGAGAAAAATAGTATCAGACATCTGGAACCTGCGCTCTTGGCCA 3717
Db 8650 GGAAGAAGTATGAGAAAAATAGTATGATTAAGATCTGGAACCAATGCGCTCTTGGCCA 8709
QY 3718 CAAGGGATCTTCTGTTAGTCAATGTCTTCAGAGTGAAGTGAACCGTCAATATCAC 3777
Db 8710 CAAGGGATCTTCTGTTAGTCAATGTCTTCAGAGTGAAGTGAACCGTCAATATCAC 8769
QY 3778 GAGCGAGCATCTGAGAAATTCATGCAACGCTGGAAGAAAAAGATCAGAGAAATTTGTGCG 3837

```
|||||
Db 8770 GAGGGAGCATCTGAGATTATCATCCGTGAGAAAAGATCAGAGGAAGTTTGTCCG 8829
Qy 3838 TAGAGAGAGATCTGTTTCCACCCGTCATGAGAAAGTGTAAAGTCCACGTTTACGA 3897
Db 8830 TAGAGAGAGATCTGTTTCCACCCGTCATGAGAAAGTGTAAAGTCCACGTTTACGA 8889
Qy 3898 TCACCTTGAAGAGACGTCTGCGGGTACATACCATGACACAGGCGCAGCCACAGCGTA 3957
Db 8890 TCACCTTGAAGAGACGTCTGCGGGTACATACCATGACACAGGCGCAGCCACAGCGTA 8949
Qy 3958 TAACTCTATCTGAGAGAGCGTCAGGGAAGTGTATTAACACCTTCTGCAAGAA 4017
Db 8950 TAACTCTATCTGAGAGAGCGTCAGGGAAGTGTATTAACACCTTCTGCAAGAA 9009
Qy 4018 CGTCACTCAAGATGTAAGTGTGCGACTACAGCAGGTAATCTGTAGCAGCGCAAGAA 4077
Db 9010 CGTCACTCAAGATGTAAGTGTGGACTACAGCAGGTAATCTGTAGCAGCGCAAGAA 9069
Qy 4078 GATGAAGGGCTGCACTAAAGCAAAAGTGATTCCTACAAAGCGCAAGGAATG 4137
Db 9070 GATGAAGGGCTGCACTAAAGCAAAAGTGATTCCTACAAAGCGCAAGGAATG 9129
Qy 4138 GGTCTTCAACTCGCCGATCTTATTAAGCACACAGACCACTCAGTGCAGAGTTAATTGCA 4197
Db 9130 GGTCTTCAACTCGCCGATCTTATTAAGCACACAGACCACTCAGTGCAGAGTTAATTGCA 9189
Qy 4198 CATTCCTATCCGCTTGACACCGACAGCTGTCCCGGTTCCGTTAGCTCACAGCCTACAGT 4257
Db 9190 CATTCCTATCCGCTTGACACCGACAGCTGTCCCGGTTCCGTTAGCTCACAGCCTACAGT 9249
Qy 4258 CAGAGAGGTTCAAAGGACATCACCCCTCACCTGACTGCAATGCGAACCAATTCGTGAC 4317
Db 9250 CAGAGAGGTTCAAAGGACATCACCCCTCACCTGACTGCAATGCGAACCAATTCGTGAC 9309
Qy 4318 AACGAGAAATTTGGGGCTGCGAGAGACAGCAACAGCAGAAATGATTACAGGGTCTACATC 4377
Db 9310 AACGAGAAATTTGGGGCTGCGAGAGACAGCAACAGCAGAAATGATTACAGGGTCTACATC 9369
Qy 4378 CAGAAATTTTCTGTGGGGCGAAGAGGCTGGAATGATGATGGAATCACTGAACCAAGT 4437
Db 9370 CAGAAATTTTCTGTGGGGCGAAGAGGCTGGAATGATGATGGAATCACTGAACCAAGT 9429
Qy 4438 CAGAGTCTGGGCCAGAGAGTGGGACCCAGCGACCCACATGATGAGCGCATGAGATCAT 4497
Db 9430 CAGAGTCTGGGCCAGAGAGTGGGACCCAGCGACCCACATGATGAGCGCATGAGATCAT 9489
Qy 4498 CATCACTATTTATCATCGGCATCCAGTCTACATGTCATTTGTGTGTGTGTGTGTGTGT 4557
Db 9490 CATCACTATTTATCATCGGCATCCAGTCTACATGTCATTTGTGTGTGTGTGTGTGTGT 9549
Qy 4558 TGCATCTCTGTGATGAGCTGATCATCAGCAGCTTGCATCGCCAAAGAAAGAAAGACTG 4617
Db 9550 TGCATCTCTGTGATGAGCTGATCGTCAAGCAGCTTGTATCCCAAAGAAAGAAAGACTG 9609
Qy 4618 CCTGACGCGCATAGCGGCTTGACCGAAGCAACGGTACCCACAGATTAAGCGGTTTGTG 4677
Db 9610 CCTGACGCGCATAGCGGCTTGACCGAAGCAACGGTACCCACAGATTAAGCGGTTTGTG 9669
Qy 4678 CTGATTCGGGCAACCAACGCTGAAACATTTGAGAAACCTTGAACCATCTGTGGTTTAA 4737
Db 9670 CTGATTCGGGCAACCAACGCTGAAACATTTGAGAAACCTTGAACCATCTGTGGTTTAA 9729
Qy 4738 CAACCAACCGTTTCTCTGGGCAAGTGTGATCTCTGCGACGCGCTTGTATTTCTGT 4797
Db 9730 CAACCAACCGTTTCTCTGGGCAAGTGTGATCTCTCTGCGACGCGCTTGTATTTCTGT 9789
Qy 4798 CCGGCTCTTTCATGCTGCATGCGCTTTTATTTGTTGTCAGAGCGCTGCGGGAAGGT 4857
Db 9790 CCGGCTCTTTCATGCTGCATGCGCTTTTATTTGTTGTCAGAGCGCTGCGGGAAGGT 9849
Qy 4858 AGAGCGCTTTCAGAGCTGCAACACTGTGCAAAATGTTCCGGGATCCCGTATTAAGCGCTT 4917
|||||

Db 9850 AGAGCGCTTTCAGAGCTGCAACACTGTGCAAAATGTTCCGGGATCCCGTATTAAGCGCTT 9909
Qy 4918 GGTGGAACGGGAGGTTAAGGCGCACTTAACCTGAGATCAGGAGTGTCTCATCGGAATT 4977
Db 9910 GGTGGAACGGGAGGTTAAGGCGCACTTAACCTGAGATCAGGAGTGTCTCATCGGAATT 9969
Qy 4978 AACACCTTCAACTAACAGAGATACGTGACCTGCAAATTCACACAGTCAATTCCTTACC 5037
Db 9970 AACACCTTCAACTAACAGAGATACGTGACCTGCAAATTCACACAGTGTCTTCTTACC 10029
Qy 5038 ACAAGTTAAATGCTGGGGTCCCTCGAGTCAAGGCACTCTCAAAAGCGGATTAACATG 5097
Db 10030 ACAAGTTAAATGCTGGGGTCCCTCGAGTCAAGGCACTCTCAAAAGCGGATTAACATG 10089
Qy 5098 CCGGTTTTTGGGGGTGTACCTTTCATGTTGGGAGGGGCAATGCTTCTGTGACAG 5157
Db 10090 CCGGTTTTTGGGGGTGTACCTTTCATGTTGGGAGGGGCAACAGTGTCTGTGACAG 10149
Qy 5158 TGAGAACACAACTGAGTGAAGGCGTACGTCGAGTTGCTTCAGACTGCATATAGATCA 5217
Db 10150 TGAGAACACAACTGAGTGAAGGCGATACGTCGAGTTGCTTCAGACTGCATATAGATCA 10209
Qy 5218 CGCAGTCGACTAAAGTTCAACAAGCTGCTTGAAAGTGGCCCTGCTATAGTATACGG 5277
Db 10210 TGCAGTCGACTAAAGTTCAACAAGCTGCTTGAAAGTGGCCCTGCTATAGTATACGG 10269
Qy 5278 CAACACACCGGCGCACTGATAGTTCATAGGCTGACGCGAGGTTCTTCAAGGGA 5337
Db 10270 CAATACCAAGCGGCGCTGATACATTCGTACACGGGCTCAACAGGTTCTTCAAGGGA 10329
Qy 5338 CCTGAAGGTATAGCAGAGGCGATATCAGCGGCTTTTCAACCTTTGACCATTAAGTCTG 5397
Db 10330 CTTGAAGGTATAGCAGAGGCGATATCAGCAGGCTTTTCAACCTTTGACCATTAAGTCTG 10389
Qy 5398 CATCAGAAAGGCTTGTATTACAATGACCTTCTGAGTATGAGACTATGAACCAAG 5457
Db 10390 CATCAGAAAGGCTTGTATTACAATGACCTTCTGAGTATGAGACTATGAACCAAG 10449
Qy 5458 AGGTTGGGGATATTCGAACATCTGCTGATGATGACAGACATAGTACCCGCACTGA 5517
Db 10450 AGGTTGGGGATATTCGAACATCTGCTGATGATGACAGACATAGTACCCGCACTGA 10509
Qy 5518 CATACGCTCTGGAAGCTTCTGTCAAGAACATCCAGTCCCTTACCCCAAGCAGTATC 5577
Db 10510 CATACGCTCTGGAAGCTTCTGTCAAGAACATCCAGTCCCTTACCCCAAGCAGTATC 10569
Qy 5578 AGGTTATGAATGTGAAGAACAACTGACAGACGACCTTCAAGAAACAGCACATTTGG 5637
Db 10570 AGGTTATGAATGTGAAGAACAACTGACAGACGACCTTCAAGAAACAGCACATTTGG 10629
Qy 5638 ATGTAAATTTGAAGTGAAGCTGCGAGGCTCTAATGTGCTTACGGGACATCCCTAT 5697
Db 10630 ATGTAAATTTGAAGTGAAGCTGCGAGGCACTAATGTGCTTATGGGACATCCCTAT 10689
Qy 5698 CTCGATTCAGATCCCTGATGACGCTTTGTGAGATCATCAGAAATCAACCAATTTTGA 5757
Db 10690 CTCGATTCAGATCCCTGATGACGCTTTGTGAGATCATCAGAAATCAACCAATTTTGA 10749
Qy 5758 AGTTAGCTGACAGTACAGACGTGATTTATTTGCGAGACTTTGTGTGTCTCTCAACAT 5817
Db 10750 AGTTAGCTGACAGTACAGACGTGATTTATTTGCGAGACTTTGTGTGTCTCTCAACAT 10809
Qy 5818 ACAGTACAAAGCTGACAGGAGGGAATGTCCAGTTCACTCCCACTCCAGACAGCTGT 5877
Db 10810 ACAGTACAAAGCTGACAGGAGGGAATGTGCAAGTTCACCTCCCACTCCAGACAGCTGT 10869
Qy 5878 TTTGAAGGAAGCGACACACATGATGATGCGGTAGGAGCAGATTAACATATTTTACAC 5937
Db 10870 TTTGAAGGAAGCGACACACATGATGATGCGCAAG---CATTAACATTAATTTTACAC 10926
Qy 5938 ATCAGGCCCAACAGCAAAATTTATAGTTTGTGCTATGCGGCAAGAGTCCACTGCAATGC 5997
Db 10927 ATCAGGCCCAACAGCAAAATTTATAGTTTGTGCTATGCGGCAAGAGACCACTGCAATGC 10986
```

```

OY 5998 TGAATGTAACACCGCGGACACATTAATTGGAGAACACATAAAGTCGACCAAGAATT 6057
DB 10987 TGAATGTAACACCGCGGACACATTAATTGGAGAACACATTAAGTCGACCAAGAATT 11046
OY 6058 CCAGCGCGGAGTTTCCAAAACATCTTGGAACTGCGCTTGGACGTTTGGGGAGCATC 6117
DB 11047 CCAGCGCGGAGTTTCCAAAACATCTTGGAACTGCGCTTGGACGTTTGGGGAGCATC 11106
OY 6118 ATCCCTCATGTTGTAGAGCTTAATAGTGTGTCTGCGAGCTTATGCTTATAACACACG 6177
DB 11107 ATCCCTCATGTTGTAGAGCTTAATAGTGTGTCTGCGAGCTTATGCTTATAACACACG 11166
OY 6178 TAGATGACGACGCGGACACTGACATAGCGGT-AAAACGATGATCTTCGAGGAAGC 6236
DB 11167 TAGATGACGACGCGGACACTGACATAGCGGTAAAACGATGATCTTCGAGGAAGC 11226
OY 6237 GTGGTGCAATATGCCAGCGCGCTTGAACACTAAACTGATGATTTCCGAGGAAGC 6296
DB 11227 GTGGTGCAATATGCCAGCGCGCTTGAACACTAAACTGATGATTTCCGAGGAAGC 11286
OY 6297 AGTGATATAAGTGTGACGCTGACATTAATGTAATACACTACATTTTAAACACT 6356
DB 11287 AGTGATATAAGTGTGACGCTGACATTAATGTAATACACTACATTTTAAACACT 11346
OY 6357 ATATCACTTTTATGAGACTCACTATGGGCTCTAATATATACACTACATTTTAA 6416
DB 11347 ATATCACTTTTATGAGACTCACTATGGGCTCTAATATATACACTACATTTTAA 11406
OY 6417 AAACACTATACACTTTATAAATCTTTTATTAATTTTCTTTTGTGTTT 6476
DB 11407 AAACACTATACACTTTATAAATCTTTTATTAATTTTCTTTTGTGTTT 11466
OY 6477 TAAATTT 6484
DB 11467 TTAATAAT 11474

```

RESULT 2

AAV33475
ID AAV33475 standard; cDNA, 11663 BP.

AAV33475;

10-DEC-1998 (first entry)

South African Arbovirus strain No. 86 cDNA clone p555 sequence.

Bone marrow cell; alphavirus; South African Arbovirus strain No. 86;
KW S.A.AR86; reverse transcriptase-PCR; nsP1; nsP2; nsP3; nsP4; capsid;
E3; E2; 6K; E1; growth hormone; growth factor; interleukin; cytokine;
chemokine; enzyme; ribozyme; antisense oligonucleotide; ss.

South african arbovirus.

Key Location/Qualifiers

215 /tag= a "This nucleotide varies from A to G with respect
FT /note= to the S.A.AR86 cDNA sequence given in AAV33473"

variation 3863 /tag= b "This nucleotide varies from C to G with respect
FT /note= to the S.A.AR86 cDNA sequence given in AAV33473"

variation 5984 /tag= c "This nucleotide varies from A to G with respect
FT /note= to the S.A.AR86 cDNA sequence given in AAV33473"

variation 9113 /tag= d "This nucleotide varies from T to C with respect
FT /note= to the S.A.AR86 cDNA sequence given in AAV33473"

variation 5947 /tag= e "This nucleotide varies from T to C with respect
FT /note= to the S.A.AR86 cDNA sequence given in AAV33473"

PN WO9836779-A2.
XX 27-AUG-1998.
PD 18-FEB-1998; 98WO-US02945.
XX 19-FEB-1997; 97US-0801263.
PR (UNNC-) UNIV NORTH CAROLINA.
PA Davis NL, Johnston RE, Simpson DA,
XX WPI, 1998-495361/42.
DR P-PDB; AAW70471, AAW70472.
XX

PT Expression of heterologous RNA in bone marrow cells - using a
recombinant alphavirus comprising a promoter operable in bone marrow
cells operably associated with heterologous RNA.

Example 5; Fig 5A-5B; 68pp; English.

The present sequence represents the South African Arbovirus strain
No. 86 (S.A.AR86) cDNA clone p555 sequence used in the method of
the invention. The invention provides a method for introducing and
expressing heterologous RNA in bone marrow cells using alphavirus
vectors. S.A.AR86 is an alphavirus whose cDNA sequence (AAV33474) was
determined from uncloned reverse transcriptase-PCR reaction fragments
amplified from the virion RNA. The S.A.AR86 cDNA clone p555 sequence
has four nucleotide variations when compared with the S.A.AR86 cDNA
sequence derived from genomic RNA. The variations, however, do not
alter the sequence of the protein encoded with respect to the
S.A.AR86 genomic RNA. The S.A.AR86 cDNA encodes nonstructural
and structural polypeptides. The nonstructural polypeptide is,
presumably, post-translationally modified into four different protein
products, namely nsP1 (AAW70460), nsP2 (AAW70461), nsP3 (AAW70466),
nsP4 (AAW70467) proteins. The structural polypeptide is, presumably,
post-translationally modified into five different protein products,
namely capsid (AAW70468), E3 (AAW70469), E2 (AAW70470), 6K (AAW70471),
E1 (AAW70472) proteins. The S.A.AR86 cDNA was used in the method of the
invention. The inventors claim the transformed bone marrow cells are
useful for expressing a protein or peptide suitable for protecting the
subject against a disease such as a microbial, bacterial, protozoal,
parasitic or viral disease. The transformed bone marrow cells are also
claimed to be useful for expressing proteins and peptides such as
hormones, growth hormones, growth factors, interleukins, cytokines,
chemokines, enzymes, ribozymes or antisense oligonucleotides.

Sequence 11663 BP; 3289 A; 3105 C; 2910 G; 2359 T; 0 other;

Query Match 37.0%; Score 2397.6; DB 19; Length 11663;
Best local Similarity 64.6%; Pred. No. 0;
Matches 3152; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;

```

OY 580 ACTGACGGTATGAACCGGAGCGTATATTTCTCATCGGAACAGCCAGTCACTTC 639
DB 5707 ACTGCTAACCGGGGTAGTGGTATTTTGTGACGACACAGCCCTGGCATCTGC 5766
OY 640 AACGAATTCAGTACGTATGTAATCAACGAACCTATATGATCGGCGCTCATG 699
DB 5767 AAAAGAGTCCGTTCTGCAGAACCACTTACAGAACCACTTGGAGGCAATGTTG 5826
OY 700 AGAAGTATGAGCCCGCGCTCGATCTCGAAAGAGGAAATGTTACAGAAAGTGC 759
DB 5827 AAAAGATCTAGCCCGGTGCTGACAGCTGAAAGAGGAAACACTCAACTCAGTAC 5886
OY 760 AATTATGCGCTCTGAAGGAATATGAAGCAGGTATCAATACGAAAGTATGATATGA 819
DB 5887 AGATGATGCCACCGAAGCCAAACAGAGTACATCTCGAAAGTATGATATGACGA 5946
OY 820 AAGCAATTACGCGGAGGAGTCAATTTCTGATTTGGGCAATATCTATCAGAAAGTGA 879
DB 5947 AAGCAATTACGAGGAGGAGTCAATTTCTGATTTGGGCAATATCTATCAGAAAGTGA 6003

```

QY	880	ATCTCTGTGAGGTGTTACAGAGTCATTAATCTCTGTACCAATCTACTCTGTCCAAACGGTAATTA	939
Db	6004	ATGACCCAGAAAGCTATTAAGATCAACTCCGMAAACCATCGTATTTCCAGCAGTGAATCCAG	6063
QY	940	ACAGGTTTAACTCTCAGAGGTCGCGGCTTAAACCGTCGACCTTAGTATTCAGAGAAAT	999
Db	6084	CGAACTACTCTGACCCCAAGTTTGTCTGTAGCTGTTTGTAAACATATCTGATGTAGAAATT	6123
QY	1000	ACCTTACAGTAGCCAGTTATTTGTATTAACAGATGAATACGATGCGTATCTTGAATGTGTGG	1059
Db	6124	ACCCGACGATGACCTTATCATGAGATCACCGAGAGTAAGATGCTTACTTGGATATGTAG	6183
QY	1060	ACGGGGCATCGTGTCTGTAGATTAAGCCATTTTTTGTCCGGCTTAACTGAGGAATGCC	1119
Db	6184	ACGGGACAGTCGCTTGGCTTAGTATCTGAACTTTTGTCCCGCCAAAGCTTGTGAAGTTAAC	6243
QY	1120	CAAAAGACATAGTATTAATTTGACGAGCAGATTAAGATTCAGCCGCTCCATCGCTTATACGA	1179
Db	6244	CGAAAGACAGAGTATTAAGAGCCCAACATCCGATGTCCGTTTCACTAGCATGACGA	6303
QY	1180	ATACATTAACAAAATGATATTTGGCTGACGTACTTAAAGGAATTCGAAACGTTACCCAAATGC	1239
Db	6304	ACAGGTTGCAAAACGTGTGCATTTGCCGGGACTTAAAGAAACCTGCAACGTCAACAAATGC	6363
QY	1240	GAGAAATTACCTGTCTTAAATTTGGCGGCGCAATTAATGTGATTTGTTCAAGAAATACGAT	1299
Db	6364	GTTGAATCTGCACACACTGGAATCTGACGGAATTCAGTTCAGTTGAATGTCTTGGAAAATATGAT	6423
QY	1300	GCAATGATGAGTACTGGGGATACCTTTGGCGATTAACCTTATCCGCTATCTACAGGAACG	1359
Db	6424	GCAATGACGATTAATTTGGAGAGATTGTGCCCGAAACCAATTAAGATCTACTGTGTTGC	6483
QY	1360	TTACGCAATATGTGACAAAGCTGAAAGGCGCGAAAGCAGCAGCATTTGTTCCGAATATCTC	1419
Db	6484	TTACGCGCATATGCTGGCCAGACTGAAGAGGCCCTTAAGCCGCCGCACTGTTCGCAAAAGACG	6543
QY	1420	ATAATCTTAAACCGTTGACGAGATACCAATGATCAATTTGTCTATGGAATCTTAAAGAG	1479
Db	6544	ATAATTTTGTCCCATTTGCAAGAGTGCCTATGATAGATTTGTCTCATGCAATGAAAAG	6603
QY	1480	ATGTCAAAAGTTACTCCCGGCAACGAACATACAGAGACCGGCTTAAGGTGACGTTATTC	1539
Db	6604	ACGTAAAGTTATCACTGCGACAGAAACACAGAAAGAACCGAAATCAAGATATC	6663
QY	1540	AGGCTGACAGATCCCTTGTCTACCGCTTACCTTTGGGGATCCATCGGAATTAAGTCCGTA	1599
Db	6664	AAGCGGAGAAACCCCTGGGACCGCTTACTATGCGGATCCACCGGGAGTTAATGTGCGCA	6723
QY	1600	GACTGAATGCGGTGCTTCTGCCAAATATCATCTCTTGGACATGTCCAGCGGAAT	1659
Db	6724	GCGTTTACAGCCGTTTGTGTACCCCAATTTCAACCGCTTGTGACATGTCCGCGGAGCACT	6783
QY	1660	TTGATGCAATTAATTTGCTGAACATTTCCACACCGGGAACCGAGTATGGAAGACGACATCG	1719
Db	6784	TTGATGCAATCAATACAGAAACCTTCAAGCAAGGACCCGATCTTGAAGACGATATCG	6843
QY	1720	CGTCGTTTGAATTAAGCGAAGCGACGCTATCGCAATTTCCGCGTTGATGATCTTGTAGG	1779
Db	6844	CCTCGTTTGAACAAAGCCAAAGCAGACGCTATGCGCTTAAACCGGCTGATGATCTTTGAGG	6903
QY	1780	ACTTAGGTGTGACCAACCGCTCTTAGATTGTGATGAGCGGCGCTTCGGCAATATCATAT	1839
Db	6904	ACCTGGGTGTGACCAACCACTACTGACTTGATGAGTGGCGCTTTGGAAGAAATATCAT	6963
QY	1840	CTGTGCACTTACTTACAGGAACGAGGTTTAAATTTGTGTCATGATGAATTCGGTATGT	1899
Db	6964	CCACCATCTGCGCAACGGGTACCCGTTTCAATTTGCGGCGCATGATGAAATCCGAATGT	7023
QY	1900	TCTTAACGCTGTTTCTCAACACACTAGTCAATATCATATGATTTGTAAGCAGAGTACTAGTG	1959
Db	7024	TCTCAACGCTTCTTGTCAACACATTTCTGAATGTGTTATGTGCGACGACAGATATTTGAGG	7083
QY	1960	AACGGTTAACCACTCAGGTGCGCGGCTCTATCGGACGATTAACATAGTGCATGTGT	2019

Db	7084	AGCGCTTAAAA	CGCTCAATGTGACAGATTTATGGGACGACACAACTTATACCGAG	7143
Qy	2020	TCGTCCTCCGACAC	CTTGATGCGGAGATGCGCCACCTTGACATGGAAGTAAAA	2079
Db	7144	TAGTATCTGACAA	AGAAATGGCTGAGAGGTGTGCGACTGGCTCAATGAGGTTAAGA	7203
Qy	2080	TTATTGATGACGTTA	TGTTATCTCAAAGCACCTTACTTGTGGGGGATTTATCTGGTGG	2139
Db	7204	TCATTGACGCGAT	CGGCGAGAGACCACTTACTTGTGGGGGATTTACTTTGCAAG	7263
Qy	2140	ACCGATPAACGGA	CAGCGCTGAGAGCGGACCGCTTAAAAAGGCTTTTAAGCTTG	2199
Db	7264	ATTGCGTTACTC	CAACAGCGTGTGCGTGGCGACCCCTTGAAAAGGCTGTTTAAGTTGG	7323
Qy	2200	GAATAACATTCG	AGTCAGTATACCCAGACTGCGACCGCGCGCGGACCTGCATGATG	2259
Db	7324	GTAACCGCTCC	CAGCGCGACATAGGACAGAGACAGAAAGCGGCTCTGCTATAGT	7383
Qy	2260	AAGCAATGCGAT	GAAACGAATTCAGAGGAGGCGGTGATGAGAGCGGTGATATCA	2319
Db	7384	AAACAAAGCGCT	GGTTTAAAGGTATTAACAGACCTTATGAGTGGCGGTGCAATC	7443
Qy	2320	GATACGAGTAT	ACTGCGAGGCTGATCATCAAGTCTGTGTCACGTTAGCCGAAGCG	2379
Db	7444	GGTATGAGT	GAGCAATCACTGCTGCGATTTGAGAACTTTGGCCAGAGCA	7503
Qy	2380	TTAAGAACTT	CAAGAGCATTAAGAGGAGCCCAATCACCCTCAAGCGCTGACTTAATAGG	2439
Db	7504	AAAGGCACTT	CAAGCCATCAGAGGGAATTAAGCATCTCTACGGGTGCTTAAATAGT	7563
Qy	2440	TGACGTAGT	AGTAA-----CAGCAGCTTACCACCGCGAGATGTTTCA	2482
Db	7564	CAGCATAGT	ATCATTTCACTAATACCAACAACACCAACCATTAATAGAGATTC	7623
Qy	2483	TACCTCAGCT	GAATCTTCCACAGTTTACCTTCAATCCGATGCGTTACCGAGATCA	2542
Db	7624	TTAACATG	TCTGGCGCGCCCTTCCAGCGCCACCTGCGATGTGAGGCGCGGAGAA	7683
Qy	2543	AAACCTCTAG	AGCGCGCTTGAGGCGCTTTCGCGCCCGCTGCGTCTCAATCGAAGAT	2602
Db	7684	GAGGCGAG	CGGCGCCCATGCTGCTGCGCAATGGGCTGTCTCCCAATCCAGCACTGA	7743
Qy	2603	CTTAGAGT	CGATAGTCAACTTGAACCTTCAACAC--GATCACTTAATCCGCGCAG	2660
Db	7744	CCACAGCG	CGTACGCTTATGTCACTTGTGAGACAGCACTTAACCTCAACCCACGCCCCAC	7803
Qy	2661	GTTCCACCG	CCCAAGAAAGAGTGTCTCTTAAAGCCAAACCTTACCTTAAAGAA	2720
Db	7804	GCCGCGCG	CGCCGACAGAGGCGCCCAAGCAACACCGAGCCGAAGAACAA	7863
Qy	2721	AGAGGAGCA	AGCCAAAGAGCAAGC---GCAAGCTTAAACAGGGAGAAAGCAACGTA	2777
Db	7864	AAACACAG	GAAGAAAGAAAGCACTCTGCAAAACCCAAACCGGAAAGAGACGCTTA	7923
Qy	2778	TGTGTATGA	ATTGAGTCCGACACATTTCCGATC---ATGCTAAGCGGCCAAGTGA	2834
Db	7924	TGGCACTTA	ATTGAGAGGCGACGACTGTTCGACGTCGCAAAAATGAGAGCGAGATGTCA	7983
Qy	2835	ATGATATG	CTGCTGTGCGAGAAAGCTGATGAAACCACTTCAAGTTGAGAGAAAA	2894
Db	7984	TGCGGCA	CGCACTGGCATGAGAGAAAGGTAAAGAAACCACTTCAAGTAAAGAACTA	8043
Qy	2895	TTGATTAAT	AGCAATTAAGCGCGCTGAAATTGAAGAAAGGCTAGATATGACCTTGAGT	2954
Db	8044	TTGACCA	CCCTGTCTATCAAAGCTCAAATTCACCAAGTGTGACATACGACTGAGAT	8103
Qy	2955	ACGCGAG	AGTTCGCCAGAACTGAATACAGACGCTGACATACCAAGCAACAAACAC	3014
Db	8104	TGCGACAG	TTCGCTCAACTGAGAGTGAAGGCTTCACTTACCAAGTGAACACCTG	8163
Qy	3015	CGGGCTT	CACTGCGACCGCGCGAGTCCAGTATGAGATGGAGATTTTACGCTAC	3074

Dh 8164 AAGGTTCTACACTGCAACGAGGCGGTGCAATATAGTGAAGGAGATTTACATCC 8223
Qy 3075 CGAGAGAGTGGGGCGGAAAGGCGACAGGGAAGCCGATCTGTGACAAAGAGGACAG 3134
Db 8224 CCCGCGAGTGAAGAGGAGAGAGAGAGTGTGCTCCGATTTGATTAATCTAGGCGGG 8283
Qy 3135 TTGTGGCTATTTGTTTAAGAGGTGCAAAATGAGGGGACGCGTACGCGCTTTCAAGTGTCA 3194
Db 8284 TTGTGCGCATAGTCTCCGAGGGGGCTGATGAGGAAACAAGAACCGCCCTTTGCGTGTCA 8343
Qy 3195 CTGGAACACGAAAGGGGTGACATTTAGGGATACCCCGAAGTTTGAACCTGTGT--- 3251
Db 8344 CCTGGAATAGCAAAAGGAGAGCAATCAAGCAACCCCGAAGGAGCAGAGAGTGTGTG 8403
Qy 3252 -----CACTAGTTACAGGCGTATGCGTCTTGATGTGCAAGTTCCCATGCGACAAAC 3305
Db 8404 CTGCAACCACTGTGTCAGGCGCATGTGTGTTGGAAGCGTAGCTTCCATGCAATCGCC 8463
Qy 3306 CACCGGTGTCTATTCACTGACGCGAAGACGAACTGACGCTGCTGAGAGAACTCG 3365
Db 8464 CGCCACATGTCTACACCCGCGAAGCATCGAGAGCTCTGACATCTCTGAAAGAAAGTGA 8523
Qy 3366 ACAATCAAATTTAGACAGCTGCTGAGAAAGCTTGAAT--GTCCATCGCCGCG 3422
Db 8524 ACCAAGAGGCTACGACACCTGCTCAAAGCATATTTGGGTGCGGATGCTCGGCAAG 8583
Qy 3423 CCAAGCAAGCACTACGATGATCTTCACTGACCAAGTCCCTACCTGGGGTTCGCCGT 3482
Db 8584 GTAAAGAAAGGCTCACTGACGATTTACCTTGACCAAGCCGTAATTTGGGACATCTCTGT 8643
Qy 3483 ATTGCAGACACTCAACGCGGTGTTTCAAGCCCAATTAATTAAGAAAGTGTGAGCAAT 3542
Db 8644 ACTGTCACTATCTGAACCGGTCTTTAGCCCGATTAAGATGACAGAGTCTGGATGTAAG 8703
Qy 3543 CTGATGATGATCGATTAGATTCAGGTCTGCGCACAATTCGGTAAATGAGGAGCA 3602
Db 8704 CGGAGCAACAACCAATAGCATACAGACTTCGCGCAAGTTGATGACCAACCAAGGAG 8763
Qy 3603 CTGCGGATGTCAACAAATTCGTTACATGCTTTTGACACCAAGCATGATCAAGAG 3662
Db 8764 CAGCAAGCTCAAAATAGTACCGCTACATGTGCTGAGAGAGATCAATCTGCAAGAG 8823
Qy 3663 ACAGATGAGAGAAATAGTATCAGACATCTGACACCTGCGTGTGCGCAACAAG 3722
Db 8824 GCACCATGATGATCAATCAAGATCAGACCTCAGAGACCGTGTGAAGGCTTTAGCTCAAG 8883
Qy 3723 GGTACTCTCTGTTAGCTCAATGTCTCTCAAGGTGACAGTGTAAACGTCAGATACAGAGCG 3782
Db 8884 GATACTTTCTCTCGCAAGTGTCTCCAGGGGAGAGCGTAAAGGTTAGCATAGCGAGTA 8943
Qy 3783 GAGCATCTGAGAAATTCATGCAACCGTGAAGAAATCAAGAGAAAGTTTGTGAGAG 3842
Db 8944 GCACCTCAGCAACGTATGACCAATGTGCGCGCAATTAACCAAAATTCGTGGAGACGG 9003
Qy 3843 AGGAGTACTTTGTTCCACCGGTGCAATGAAAGCTGTAAAGTGCACAGTTTACATCACT 3902
Db 9004 AAAAATATGACTTACTCTCCGTTACGATGAGAAAGTTCTTTCACAGTGTGACCGTTC 9063
Qy 3903 TGAAGGAGACGTCTGCGGGTACATTAACATGACAGGCGGACCAACGCGTAAAGT 3962
Db 9064 TGAAGAAACAACCGCGCTACATCACTATGCAAGGCGGAGCGGACGCTATTAAT 9123
Qy 3963 CCTATCTGAGAGAACGCTCAGCGCAAGTGTACATTAACCACTTCTGCGAAGAACGTCA 4022
Db 9124 CCTATCTGAGAAATCATCAAGGAAAGTTTACGCAAGGCAACATCTCGGAGAAAGACATTA 9183
Qy 4023 CCTACGAATGTAAAGTGTGCGACTACAGCAGAGTATCGTGACGCGGAAGAGATGA 4082
Db 9184 CGTATGAGATGCAAGTGTGCGGCAATTAAGACCGGAACGTTTACGACCGGTACGAAATCA 9243
Qy 4083 ACGGCTGACATTAAGCAAAAGTGCATTTGCTTACAGAGCGCAACCAAGAAATGGGTCT 4142
Db 9244 CGGGCTGACCGGCAATCAAGAGTGTGTGCTTATTAAGGCAACCAAGAAAGTGGGTCT 9303

Qy 4143 TCAACTGCGCGGATCTTATTTAGGACACAGACCACTAGTGAAGTAAATTTGACATTC 4202
Db 9304 TCAACTGCGCGGATCTTATTTAGGACACAGACCACTAGTGAAGTAAATTTGACATTC 9363
Qy 4203 CATTCGCTTGAACCGACAGTGTGCGCGTTCGTTAGTCAACGCGCTACAGTCAAG 4262
Db 9364 CTTCACAGCTATCCCGAGTACTGTAGATGTCCCTGTGTCCCAAGCGCGCAAGTGAAC 9423
Qy 4263 AGTGTGAAGGATCACTCTCACTGATGCAATGCGACCAATTTGCTGACAAAG 4322
Db 9424 ACGGCTTTAAACATATGAGTGTGCAACAGACCAATGATGATGATGCTACACACA 9483
Qy 4323 GAAATTTGGGCTGAGACAGCAACAGAGATGATTAAGGAGTCTACATCAGAG 4382
Db 9484 GGAAGTGGGCAACCGGCAACCACTGAATGATGATGAGAAACAGGTTAGAA 9543
Qy 4383 ATTTTCTGTGGGCGAAGAGGCTGAGTACGATGAGGTTAACCATGACAGTCAAG 4442
Db 9544 ACTTCACGCTGACCGAGATGAGCTGGAATACATATGAGGCAATCAAGAACAGTAAAG 9603
Qy 4443 TCTGGGCCAGAGTGTGACACAGGACCAACATGATGAGCGCGATGATCATATCC 4502
Db 9604 TCTATGCCCAAGATCTGACACAGAGACCTCAAGATGCGCACACGAAATAGTACAG 9663
Qy 4503 ACTATATCATGCGGATCAGTCTCACTGTCAATTTGCTGTGTGTGCTGTGCTGCT 4562
Db 9664 ATTTACTATATGCGGATCTGTGTGATGACATCTTAACCGCTGATCACTGCTGTGCGA 9723
Qy 4563 TCTGTGAGGACATGATCATCAGAGCTTGCATCGCCAAAGCAAGAGAGATGCTGTA 4622
Db 9724 TGATGATTTGGGTAAGTGTGACAGATTAATGTGCTGTAAGCGCGCGGTAGTGTGTA 9783
Qy 4623 CGCCATACGCGCTTGTGACCGAAGCAAGTAAACCAAGCATTAAGCGGTTTGTGTGCA 4682
Db 9784 CGCCATATGCGCTGTGCGCCCAATGCGTATTCACATCTGCTGTGCTGTGCTGTG 9843
Qy 4683 TTGGCCCAACCAAGCTGAACATTTGAGAACTTTGAACCATCTGTGTTTAAACAAC 4742
Db 9844 TTAGTGTGCTGAATGTGAACATTTACCGAGACATGATGATTTATGATGCAACAGCC 9903
Qy 4743 AACGTTTCTGTGGGACAGTGTGATTCCTCTGCGAGCGCTTGTATTTGTTCCGCT 4802
Db 9904 AGCGTCTTCTGTGGTCCAGCTGTGTATCTCTGCGCGCTGTGTGCTTATAGCGCT 9963
Qy 4803 GCTTTATGCTGATGCTTTTATTTATTTGTTGCAAGCGCTGTGCTGCGGAAAGTGAAG 4862
Db 9964 GTTGTCTATGCTGTGCTGCTTTTATTTAGTGTGCGCGGCTTACCTGTGCGAAGTGAAG 10023
Qy 4863 CTTTGAACATGCAACCACTGTGCAAAATGTTCCGGGATCCCGTAAAGCGTGTGCTG 4922
Db 10024 CTTAGGAACATGCAACCACTGTGCAAAATGTTCCAGATTAAGCGATCTGTTG 10083
Qy 4923 AACGCGAGGTTTACGCGCACTTAACCTGAGATCAAGTGTCTCATCGGAATTAAC 4982
Db 10084 AAAGGCGAGGTTAGCGCGCTCAATTTGAGATTTACTGTATGTCTCTGAGAGTTTGC 10143
Qy 4983 CTTCACTAACAGAGTGTGACCTGCAAAATTTCAACAGTCAATTTCTTACCAAG 5042
Db 10144 CTTCCACCAACCAAGATGATTAATCTGCAAAATTTCAACAGTGTGTCTCTCCCTTAAG 10203
Qy 5043 TTAATGTGTGGGCTCTCTGAGTGAAGAGATCTTCAAAAGCGGATTAACATGCGCG 5102
Db 10204 TCAAGTGTGTGGCTCTCTGAAATGTCAAGCGCGCTCAAGCAGATTAATCTGCAAG 10263
Qy 5103 TTTTGTGGGCTGTATCCTTTCAATGTGAGGAGCGCAATGCTTGTGACAGTGA 5162
Db 10264 TCTTTGAGGGGTGTACCTCTTCAATGTGAGGAGAGCAATGTTTGTGCAAGTGA 10323
Qy 5163 AACACCAATGATGAGGCGTGAAGTGTGCTCAAGACTGACATTAATGATCAAG 5222
Db 10324 AACACCAATGATGAGGCGTGAAGTGTGCTCAAGACTGACATTAATGATGAGTGAAG 10383

QY 5223 TCGCACTAAAAGTTCACACAGCTCTCTGAAAGTCGCGCTCGTATAGTATACGGCAACA 5282
 Db 10384 AGGGGATTTAAGTGTGATCTGCGCGATGAAAGTGAAGCTGCTATAGTATACGGGAACA 10443
 QY 5283 CCACCGCGCACCTGGATACGTTTGTCAAATGCGCTACCGCCAGGTTCTCACGGGACCTGA 5342
 Db 10444 CTACCAAGTTTCTAGATGTGATGTAACGAGTACACACGAGAACGCTTAAGACCTGA 10503
 QY 5343 AGGTCATAGCAGGGCCGATATACAGCCGTTTTCACCTTTGACCTATAGTCTGATCA 5402
 Db 10504 AAGTCATAGCTGAGCAATTTTCAGCATTTGTATACCAATTCACAAAGGTCTTTATCA 10563
 QY 5403 GAAAGGGCTTGTATACACTACGACTTCTCTAGATAGAGTATGAAACGAGAGCGT 5462
 Db 10564 ATCCGGCTGTGTGATACACTTATGACTTTCCGGAATAGCGAGCATGAACGAGAGCGT 10623
 QY 5463 TCGGCGATTTCAAGCATCTCTGCTTGTATGCTACAGACATAGTACCGGACCTGACATAC 5522
 Db 10624 TTGGAGACATTCAGATCACTCTTGAAGTACCAAGACCTCATCGCCAGCAGACATTA 10683
 QY 5523 GGGCTGGAAGGCTTCTGTGAAAGCAATCCAGTCCCTTACACCGAAGCATACAGGT 5582
 Db 10684 GGGTACTCAAGCTTCCGCAAGAACGTGATGCTCCGTACACGAGCGCATCTGGAT 10743
 QY 5583 ATGAAATGTGGAAGAACACTCAGAGACGACCTGCAAGAAACAGACCATTTGGATGTA 5642
 Db 10744 TCGAGATGTGGAAGAACACTCAGAGCGACCTGCAAGAAACGACCTTTGGTGTCA 10803
 QY 5643 AATTTGAATGAGACCTCTGCGAGCGTCTAACTGTCTTACGCGACATCTTATCTGCA 5702
 Db 10804 AGATTTGAGTCAATCCGCTTGAGACGGGTGACTGCTCATACGGAACATCTCCATTCTA 10863
 QY 5703 TTGACATCCCTGATGACGCTTTTGTGATCATAGATCAATCAACCAATTTTAAAGTGA 5762
 Db 10864 TTGACATCCGAGACGCTCTTATACAGACATAGACACGCTGCTTCAACAGTCA 10923
 QY 5763 GCTGACAGTATGACGACTGATTTATTTCTGAGACTTTGTGTCTTCTTAACATTACAGT 5822
 Db 10924 AATGTGATGTGAGAGTGTGACTTATTCAGGACTTCGAGGAGATGGCTTCCCTGAGT 10983
 QY 5823 ACAAGCTGACAGGAGGAGGACATTTGCAAGTTCATCTCCATTCACAGACAGCTTTTGA 5882
 Db 10984 ATGTATCCGACCGGAGAGGACATTTGCAAGTTCATCTCCATTCACAGACAGCTTTCC 11043
 QY 5883 AGGAAGGACCAACATGTGCTGCTGAGGACATTAACCTATTTTATGCAATCGA 5942
 Db 11044 AAGAGTGCAGATTCATGCTCTGAGAAAGAGGCGTACAGTCACTTCAGCACCGCGA 11103
 QY 5943 GCCCACAAGCAAAATTTATAGTTTGTGCTATGCGGCAAGAAATGCACTGCAATCTGAAT 6002
 Db 11104 GCCCACAAGGCAATTCATTTGATGCTGTGTGTGTAAGAGACAAATGCAATGCAAGAT 11163
 QY 6003 GTAAACCAACCGGCGGACCAATTAATGAGAACCAATTAAGTGCACCAAGAAATTCAGG 6062
 Db 11164 GCAAAACACACAGCTGATCATCTGAGACACCCGCAAAATAAGCAAAATTTCCAAG 11223
 QY 6063 CGGAGATTTCCAAACATCTTGAACCTGCGCTTGAACCTTTGGGGAGACATATCCC 6122
 Db 11224 CGGCAATCTCAAAACATCTTGAAGTGTGCTGTGTGCTTTTGGCGGCGCTCTGTGC 11283
 QY 6123 TCATTTGTGTAGGACTTATAGTGTGTGCTGACCTGATGCTTATAACACAGTGAAT 6182
 Db 11284 TATTAATTAATAGCATTAATGATTTTGTGCTGACAGATGCTGATACAGACAGAAAT 11343
 QY 6183 GACTGAGCGGCGGACACTGA-CATAGCGGTAAACCTGATGACTTCCGAGGAGCGTGT 6241
 Db 11344 GACCGGTACGCGCCCAATGACCCGACAGCAAACTCGATGACTTCCGAGGAGACTGATGT 11403
 QY 6242 GCATTAATGCCACGGCGCGCT-----TGACACT 6288
 Db 11404 GCATTAATGCATCAGCTGTGTATTAATGATCCCGCTTACCGCGGCAATATAGCAACACC 11463
 QY 6289 AAAACTGATGTATTTCCGAGGAAGCACAGTCAATATGCTGTGCAAGTGT 6318

Db 11464 AAAACTGATGATTTTCCGAGGAAGCGAGTGCATATGCTGCGCAAGTGT 11513
 RESULT 3
 AAV3473
 ID AAV3473 standard; cDNA; 11663 BP.
 AC AAV3473;
 XX
 DT 10-DEC-1998 (first entry)
 XX
 DE South African Arbovirus strain No. 86 cDNA sequence.
 XX
 KW Bone marrow cell; alphavirus; South African Arbovirus strain No. 86;
 KW S.A.ARB6; reverse transcriptase-PCR; nsP1; nsP2; nsP3; capsid;
 KW E3; E2; 6K; E1; growth hormone; growth factor; interleukin; cytokine;
 KW chemokine; enzyme; ribozyme; antisense oligonucleotide; ss.
 XX
 OS South african arbovirus
 XX
 FH Key Location/Qualifiers
 FH 5'UTR 1..59
 FT /*tag= a
 FT CDS 60..7662
 FT /*tag= b
 FT /product= "Non-structural polypeptide"
 FT CDS 7608..11345
 FT /*tag= c
 FT /product= "Structural polypeptide"
 FT 11346..11663
 FT /*tag= d
 FT 3'UTR
 XX
 PN W09836779-A2.
 XX
 PD 27-AUG-1998.
 XX
 PF 18-FEB-1998; 98WC-US02945.
 XX
 PR 19-FEB-1997; 97US-0801263.
 XX
 PA (UNNC-) UNIV NORTH CAROLINA.
 XX
 PI Davis NL, Johnston RE, Simpson DA.
 XX
 DR WPI, 1998-495361/42.
 DR P-PSDB; AAW70460, AAW70461, AAW70466, AAW70467, AAW70468, AAW70469,
 DR AAW70470, AAW70471, AAW70472.
 XX
 PT Expression of heterologous RNA in bone marrow cells - using a
 PT recombinant alphavirus comprising a promoter operable in bone marrow
 PT cells operably associated with heterologous RNA.
 XX
 PS Example 2; Fig 1A-1C; 68pp; English.
 XX
 CC The invention provides a method for introducing and expressing
 CC heterologous RNA in bone marrow cells using alphavirus vectors.
 CC The South African Arbovirus strain No. 86 (S.A.ARB6) is an
 CC alphavirus whose cDNA sequence was determined from uncioned reverse
 CC transcriptase-PCR reaction fragments amplified from the virion RNA.
 CC The S.A.ARB6 cDNA encodes nonstructural and structural polypeptides.
 CC The nonstructural polypeptide is, presumably, post-translationally
 CC modified into four different protein products, namely nsP1 (AAW70460),
 CC nsP2 (AAW70461), nsP3 (AAW70466), nsP4 (AAW70467) proteins. The
 CC structural polypeptide is, presumably, post-translationally modified into
 CC five different protein products, namely capsid (AAW70468), E3
 CC (AAW70469), E2 (AAW70470), 6K (AAW70471), E1 (AAW70472) proteins. The
 CC S.A.ARB6 cDNA was used in the method of the invention. The inventors
 CC claim the transformed bone marrow cells are useful for expressing a
 CC protein or peptide suitable for protecting the subject against a disease
 CC such as a microbial, bacterial, protozoal, parasitic or viral disease.
 CC The transformed bone marrow cells are also claimed to be useful for
 CC expressing proteins and peptides such as hormones, growth hormones,

CC growth factors, interleukins, cytokines, chemokines, enzymes, ribozymes
CC or antisense oligonucleotides.

XX Sequence 11663 BP; 3291 A; 3105 C; 2907 G; 2360 T; 0 other;

Query Match 36.9%; Score 2396; DB 19; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

```
QY 580 ACTGCGATGTAAGGAGGCTATTTCTCTCGGAAACAGGCCAAGTCACTTC 639
DB 5707 ACTGCTAACCGGGGTAGGTGGTACATATTTTCAGACGACAGGCCCTGGGCACTTGC 5766
QY 640 AACGAAATTCAGTACGTCAATGTAACTACAGAACTATATTTGGATGGGCGCTCAG 699
DB 5767 AAAAGAAATCGCTTGTGCAACGACGCTTACGAAACCACTTGAAGGCAATGTCTCG 5826
QY 700 AGAAGTATTAAGCCCGCGCTCGATCTCGAAAGAGAAATGTTACAGAAAGAACTGC 759
DB 5827 AAAAGATTAAGCCCGCGCTCGACAGTCAAGAGAAAGAAAGCTCAAACTCAGTAC 5886
QY 760 AATTATGCGCTCTGAAGAAATTAAGACGATCAATCAAGAAATGAATAATGA 819
DB 5887 AGATGATCCCAACGAGCCAAACAAAGACAGTACAGTCTGAAAGATGAACAG 5946
QY 820 AAGCAATTACAGCGGAGCTCATTTCTGATGGGACATATCTATCATCAGAAATGA 879
DB 5947 AAGCATACCACTAGACGACGCTTTCAAGGCTACGACTGATTAATCTGCGACAG--- 6003
QY 880 ATCTCTGAGGTTTACAGAGTCAATTTCTGTACCAATCTACTCTGCAACGATATTA 939
DB 6004 ATACGCGCAAGTGTATTAAGATCACTTACCGAAACATCTGATTCAGAGTGTACAG 6063
QY 940 ACAGGTTTACATCTGCAAGAGTGGGTTAAACCTGCACTTATTCAGAGATTT 999
DB 6064 CGAAGTCTGACCCCAAGTTTCTGTAGCTGTTTGAACAATCTGATGATGAATTT 6123
QY 1000 ACCCTACAGTACCGAGTATTGTATTAACAGATGAATGATGATGATGATGATG 1059
DB 6124 ACCGACGCTGACATCTTATGACATCAGACGATGATGATGATGATGATGATG 6183
QY 1060 ACGGCGATCTGCTGTGTAGTACAGGCACTTTTTCGCGGCTAACTGAGAACTAC 1119
DB 6184 ACGGACAGTGGCTGTGCTAGTACTGCAACTTTTTCGCGGCTAACTGAGATTTAC 6243
QY 1120 CAAGAAACATGATGATTTTGAAGCAAGATTAAGTCAAGCGTCCATGCTTATGCA 1179
DB 6244 CGAAAGACAGAGATTAAGAGCCCAACATCCGAGTGGGTTCTCATCAGGATGACAG 6303
QY 1180 ATACATTACAAATGTATTTGGCTGAGCTAATAAAGAAATTTGCAACGTTACCCAAATGC 1239
DB 6304 ACACTTTCAAACGCTGCTCATTTGCGGACATTAAGAAATGCAACGCTCACAAATGC 6363
QY 1240 GAGATTTACCTGTCTTATGATTCGGGCACTTAAATGTTGTTCAAGAAATTCGAT 1299
DB 6364 GTGAATCTCCAACTAGTACCTAGGACATTTCAAGTTGATGCTTTGAAATATGAT 6423
QY 1300 GCATGATGATGATCTGGGATACCTTTCCGATTAACCTTATGCTTAACTACAGAAAG 1359
DB 6424 GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6483
QY 1360 TTACGATATATGTCAGAAAGGCGGAAAGGCGGAAAGGACGACGATGTTTGGAAATATC 1419
DB 6484 TTACGATATATGTCAGAAAGGCGGAAAGGCGGACGACGATGTTTGGAAAGATC 6543
QY 1420 ATAAATCTAAACCGTTGACAGAGATACCAATGATCAATTCGTCATGATCTTAAAGAG 1479
DB 6544 ATAAATTCGCTCCATTTGAAGAAAGTCCATGATGATGATGATGATGATGATGATG 6603
QY 1480 ATGTAAAGTTATCTCCGACAGAAACATACAGAGAGGCGCTTAAAGTCAAGTTATTC 1539
DB 6604 ACGTAAAGTTATCACTGCGACAGAAACACAGAGAAAGACCGAAAGTCAAGTATATC 6663
```

```
QY 1540 AGGCTGAGATCCCTTGTCTACCGCTTACCTTTGCGGATCATCGGAAATTAATGCCGA 1599
DB 6664 AAGCGGAGAAACCTTGGCGACCGCTTATCATATGCGGATCCACCGGAGTTAGTCCCA 6723
QY 1600 GACTGATGCGGCTTGTGCGAAATATCATCTCTTTGACATGTCAGCGGAATTT 1659
DB 6724 GCTTACAGCGCTTTTGTGTCACCAATTCACAGCTCTTTGACATGTCGCGGAGACT 6783
QY 1660 TTGAATGATATTTGTAATTTCCACCGCGGACCGGATTTGAAACCGACATCG 1719
DB 6784 TTGAATGATATATGCAAGAAACATTAAGAGGATGACCGGATCTGGAAGCGATATCG 6843
QY 1720 CGTCTTTGATTAAGGCAAGACGACGCTATGCGCATTTGCGCGTGTATGATTCCTTGAG 1779
DB 6844 CCTGTTGCAAAAGCAAGCAACGATATGCGGTTAACCGGCTGATGATCTTTGGAAG 6903
QY 1780 ACTTAAAGTGTGACCAACCGCTCTTAATTTGATTAAGAGCGGCTTGGCAATATCAT 1839
DB 6904 ACTTGGGTGTGACCAACCACTACTGATGATGATGATGATGATGATGATGATGAT 6963
QY 1840 CTGTGACCTTACCAAGGAAACGAGTTAAATTTGTCATGATGAATCCGATATCT 1899
DB 6964 CCACCATCTGCGCAGGATCCGCTTCAATTTGCGGCGATGATGAATTCGGAATGT 7023
QY 1900 TCTTAAAGCTTTGTCAACACATGATCAATATCATATGATGATGATGATGATGATG 1959
DB 7024 TCTTCAAGCTTTTGTCAACACATGATGATGATGATGATGATGATGATGATGATG 7083
QY 1960 AAGGTTAACAGGTCAGCGTGCAGGCTCTATTCGCGGACGATTAATGATGATGATG 2019
DB 7084 AAGGCTTAAACGCTTCAATGTCAGATTTATCGGAGCAACATTTATACAGGAG 7143
QY 2020 TCGTCTCCGACCTTGAATGCGGAGATGCGGCACTTGGCTGAACATGGAATGA 2079
DB 7144 TAGTATCTGAAGAAATGCTGAGAGTGTGCACTGCTCAACATGAGATGA 7203
QY 2080 TTATTTAGTACGATTTATGTTATCAAGACGCTTCTGTGCGGAGATTTATCTGATG 2139
DB 7204 TCAATTAAGCGATGATCGGCGAGAGCACTTACTTCTGCGGTGATGATCTTTGCAAG 7263
QY 2140 ACCAGATTAACGAGCAAGGCTGCAAGTTCGAGACCTCTAATAAGCTTTTAAAGCTTG 2199
DB 7264 ATTCGTTACTTCCACAGCGTGTGCGGTGCGGAGCCCTTGAATAAGCTTTTAAAGTGG 7323
QY 2200 GAAATCAATTCGAGTATATACCAAGCTGCGACCGCGCGGCACTGATGATG 2259
DB 7324 GTAACCGCTTCCAGCGACGATGAGCAAGACGAAAGCGCGCTCTGCTAGATG 7383
QY 2260 AAGCAATGCGATGGAACGAATTTGGAATTTAGGACGATTTGTAAGGCGGTAGAATGCA 2319
DB 7384 AAGCAATGCGATGGAATTTGGAATTTAGGACGATTTGTAAGGCGGTAGAATGCA 2319
QY 2320 GATACGATATATCTGCGAGGCTGATCATCACTCTCTGTCACGTTTACCGGAAAGCG 2379
DB 7444 GGTATGATGATGACAACTACATCACTGCTGCTGCGATGATGAACTTTTCCAGAGCA 7503
QY 2380 TTAAGAACTTCAAGACATTAAGAGAGGACCAATCACTTACGCTGACCTTAAATG 2439
DB 7504 AAGCAATTTCAAGCAATCAAGAGGGAATTAAGCAATCTTACGCTGCTTAAATG 7563
QY 2440 TGACGTATGTA-----CAGGACCTTACCAACCGGAGATGTTTCA 2482
DB 7564 CAGATATGATATTTCACTGATTAATTAACCAACCAACCACTGATATGAGATTTCT 7623
QY 2483 TACCTCAGCTGAACCTTTCACACAGTTTACCTTAAATTCGATGCTTACCGAGATCCA 2542
DB 7624 TTAACATGCTGCGCGCGCGCTTCCAGCGCCCACTGATGATGAGCGCGGAGAA 7683
QY 2543 AACCTCTTAGGCGCGCTGAGGCGCTTGGCGCGCGCTGCTGCTCAATTAAGAGAT 2602
DB 7684 GAGGACAGCGCGCGCTGAGGCGCTTGGCGCGCGCTTGGCGCGCTTGGCGCGCTG 7743
QY 2603 CTTAGAGGTGATGATCACTTGAATTTCAAAAC--GATCACTTAATCCGCGCGAG 2660
```

D 7744 CCAAGCCGTCAGTGCCTTACTCATTTGGACAGGCACTAGACCTCAAAACCCACGCCAC 7803
Q 2661 GTCCACGGCCAAAGAGAGAGAGTGTCTCTAAAGCCAAAACCTATCTACGCTTAAAAA 2720
D 7804 GCCCGCGCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7863
Q 2721 AGAAGCAGCAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2777
D 7864 AAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7923
Q 2778 TGTGTATGAAGTTGAGTGCAGACAGACATTTCCGATC---ATGCTGAACGCCAAGTGA 2834
D 7924 TGGACATTAAGTTGAGAGCCAGACAGCTGTTGACAGTCAAAAATGAGAGAGAGATGCA 7983
Q 2835 ATGATATGCTGCTGCTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2894
D 7984 TCGGGCAGCAGCTGCGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8043
Q 2895 TTGATTAATGAGCAATTAAGCGCGCGTGAATTAAGAGAGAGAGAGAGAGAGAGAGAG 2954
D 8044 TTGACACCTGCTGCTATCAAAAGCTCAAAATTCACAGAGTGTCAAGCATACAGATGAG 8103
Q 2955 ACGGCGACGTTCCCGCAGAGCATGAATCAGACAGCTGCAGTACAGACAGAGAGAGAG 3014
D 8104 TCGACAGTTCGCGGTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8163
Q 3015 CGGCGCTTACACTGCGACAGCGCGCAGCTGCAGATGAGAGAGAGAGAGAGAGAGAG 3074
D 8164 AAGGTTCTACAACTGCGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8223
Q 3075 CGAAGAGAGTGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3134
D 8224 CCGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8283
Q 3135 TTGTGCTATTTGTTCTAGAGAGTCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 3194
D 8284 TTGTGCGGATGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8343
Q 3195 CTTGGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3251
D 8344 CTTGGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8403
Q 3252 -----CACTAGTTACAGCGCTATGCGCTTGGAAATGCAAGTCTCCATGCGACAAC 3305
D 8404 CTGACACCTGCTACAGGCAATGCTTGTCTTGGAAAGTGAAGCTTCCATGCAATGCC 8463
Q 3306 CACCGCTGCTATTTCACTGACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
D 8464 CGCCCATGCTTACACCGCGGAGACATCAGAGAGCTCTGACATCTCGAAGAGAGAGAG 8523
Q 3366 ACAATCCAAATTAAGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
D 8524 ACCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8583
Q 3423 CCAAGAGAGAGATTAACGATGATTTCACTGACAGAGAGAGAGAGAGAGAGAGAGAG 3482
D 8584 GTAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8643
Q 3483 ATTGACAGACCTCAACGCGCTGTTTCAAGCCCAATTAATTAAGAGAGAGAGAGAGAG 3542
D 8644 ACTGTACCACTATGTAACCGTGTCTTACCGGATTAAGATGAGAGAGAGAGAGAGAG 8703
Q 3543 CTGATGATGATTCGATTAAGATTCAGAGTCTCGGCAATTCGGCTACATGAGAGAGCA 3602
D 8704 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8763
Q 3603 CTGAGGATGATCAACAAATTCGGTATGATGCTTTTGAAGAGAGAGAGAGAGAGAG 3662
D 8764 CAGAGAGTCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8823
Q 3663 AACAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3722

D 8824 GCACATGATGACATCAAGATGACAGACCTCAGAGCCGTGTAGAGAGCTTACAGAG 8883
Q 3723 GATATCTTCCCTTTAGTCAATGTCCTCCAGAGTCAAGTGAATCCGTACAGATCAAGAG 3782
D 8884 GATATCTTCTCCCGAGAGTGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8943
Q 3783 GACATCTGAGAAATTCATGACAGCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3842
D 8944 GCACATGAGAGAGATGACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9003
Q 3843 AGAGATGATTTTCCACCCGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3902
D 9004 AAAAATATGACATACCTCCGTTCAAGGATGAAGATTCCTTCCAGAGTGTAGAGAGCT 9063
Q 3903 TGAAGAGAGAGTGTCCCGGATGATTAACATGACAGAGAGAGAGAGAGAGAGAGAGAG 3962
D 9064 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9123
Q 3963 CTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4022
D 9124 CTTATCTGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9183
Q 4023 CTTAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4082
D 9184 CTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9243
Q 4083 ACGGCTGACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4142
D 9244 CCGGCTGACAGCAGATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9303
Q 4143 TCAATCGCGAGATCTTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4202
D 9304 TCAATCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9363
Q 4203 CATTCGCTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4262
D 9364 CTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9423
Q 4263 AGTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4322
D 9424 ACGGCTTAAACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9483
Q 4323 GAAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4382
D 9484 GAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9543
Q 4383 ATTTTCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4442
D 9544 ACTTCAACGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9603
Q 4443 TCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4502
D 9604 TCTATGCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9663
Q 4503 ACTATATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4562
D 9664 ATTAATATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9723
Q 4563 TCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4622
D 9724 TGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9783
Q 4623 CGGCTATCGGCTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4682
D 9784 CGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9843
Q 4683 TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4742
D 9844 TTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9903
Q 4743 AACGTTTCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4802
D 9904 AGCGTTTCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9963

Qy	4803	GCCTTTCATGCTGCATGCACTCTTTTATTTATGGTGCAGGCGCTGCGCTGGGGAAAGTGAAG	4862
Db	9964	GTTCTCATGCTGCTCTCTCTTTTATGTGGTGGCGGCGCTTACCTGGGCAAGTGAAGC	10023
Qy	4863	CCCTTCGAAACATGCGAACCACTGTGCGCAAAATGTTCCGGGGAATCCCGTATTAAGCGTTGGTGC	4922
Db	10024	CCTACGAAACATGGGACCACTGTTCCAAATGTGCCACAGATACCGTATTAAGCACTTGTTC	10083
Qy	4923	AACCGGCAAGTTACGCGCCACTTAACCTGGAAGTACAGGTGCTCATCGGAATTAAAC	4982
Db	10084	AAAGGGCAGGGTACGCGCCCGCTCAATTGGAGATTACTGTCAATGCTCGAGAGTTTGC	10143
Qy	4983	CTTCAACTTAACAGGAATACGTGACCTTGCAAAATTTCCACACAGCATTTCTTACCACAG	5042
Db	10144	CTTTCACCAACCAAGATTAACATTACCGCAAAATTCACACTGTGTGTCCTCTCCCTTAAG	10203
Qy	5043	TTAAATCTGCGGGTCCCTCGAGTGCAGAGCATCTCTCAAGGGGGAATTACACTGCGCG	5102
Db	10204	TCAGATCTGCGGTCCTTGGAAATGTAGCCGCGCTCAGCGAGCATATCTGTCAAG	10263
Qy	5103	TTTTTGGCGGTGTGTACCTTTCAATGTGGGAGCGCACATGCTTGTGACAGTGA	5162
Db	10264	TCCTTGGAGGGGTGTACCCCTTCATGTGGGAGAGACAAATGTTTTTGGACAGTGA	10323
Qy	5163	ACACACAACTGAGTGAAGCGGTAGTGCAGATTCCGTCACAGCTGACATTATGATCACGAG	5222
Db	10324	ACACCCGAGATGAGTGAAGCGGTAGTCGAATTTGTCAGTAGTTGCGGACGTGACACGCGC	10383
Qy	5223	TCGCACATAAAGTTCAACACAGCTGCTGAAAGTCGGGCGGTATGATTAACGGCAAC	5282
Db	10384	AGGCGATTAAGTGCATATCTGCCGATGAAGAGACCTGCTATATGTATCGGGAAC	10443
Qy	5283	CCAACGGGCACTGTGATACCTTTGTCAATGGCGCTCAAGCAGGTTCTCTACGGGACCTGA	5342
Db	10444	CTACCACTTCTCTAGATGTGTAGTGAACGAGTCAACCGAGAACGTCTTAAAGACTGA	10503
Qy	5343	AGGTCAATGACAGGCGGATATCAGCCGCTTTTCAACCCTTGAACCAATAGTGTGATCA	5402
Db	10504	AAGTCATAGCTGAGCAAAATTTACAGACTTGTTCACCATTCGATCAAGAGTGTATCA	10563
Qy	5403	GAAAGGGGCTTGTTTAACTAAGACTTCCCTGATAGTAGAGTATGAACAAGAGCGT	5462
Db	10564	ATCGCGGCTGTGTATCAACTATGACTTTCGGAATTCGAGGAGTGAACACAGAGCGT	10623
Qy	5463	TCGGCGATTTATCAAGCACTCTCGCTTGATCTACAGACATATGAGCCGCACTGACATAC	5522
Db	10624	TTGAGACATTTCAAGTACTCTCTTGACTGACAAAGACTCATGCGCAGACAGACATTA	10683
Qy	5523	GGCTGCTGAAGCCTTCTGTCAAGAACATCCAGTCCCTTACACCCAGGAGTACAGGT	5582
Db	10684	GGCTACTCAAGCCTTCGCGCAAGAACGTGATGCCGTATCAGGAGCGGCGATCTGAT	10743
Qy	5583	ATGAAATGTGGAAGAACAACTCAGAGACCCCTGCAAGAAACAGACCAATTTGGATGA	5642
Db	10744	TCGAGAGTGTGAAAACAACTCAGGCGGCCCATCGAGGAACCGCCCTTTTGGGTGCA	10803
Qy	5643	AAATTGAAGTGAAGCCTCTGCGAGCGCTTACCTGTGCTTAACGGGACATCCCTATCTGCA	5702
Db	10804	AGATTGCAATCTCCCTTCTGAGCGGTGAGCTGTATACGGGAACATTTCCATTTCTA	10863
Qy	5703	TTGACATCCCTGATGCACTTTTGTGAGATCATCAGAAATCCCAAAATTTTGAAGTTA	5762
Db	10864	TTGACATCCCGAAGCGTGCCTTTATCAGGACATCAATGACCACTGCTCAACAGTCA	10923
Qy	5763	GCTGCACAGTACGACGTGATTTATTTCTGCAACCTTTGGTGTCTCTAACATTAAGT	5822
Db	10924	AATGTGATGTCAGTGATGCACTTATTCAGCGGACTTCGAGGGAGTGGCTACCTGCAGT	10983
Qy	5823	ACAAAGCTGACAGGGAAGGACATTGTCAAGTTCACTCCACTCAGACAGCGCTTTTGA	5882
Db	10984	ATGTATTCGACCGGAAAGCAATGCCCTGTATCACTTGTGACATCAAGCAACCCCTTC	11043

QY	5883	AGGAAGCGACACACATGTGACGTCCGTTGGGAGCAATACATTAATTTTACATCA	5942
Db	11044	AAGAGTGAAGTTTCATGTGCTTGAGAAAGAGCGGTGACAGTACACTTCACGACCGCA	11103
QY	5943	GCCCAACAGCAATTTTATATGTTTGGCTATGGCGCAGAAGTCCACTTGCAATGCTGAAT	6002
Db	11104	GCCCAACAGGGAATTCATATGTAATCCCTGTGTGTGAAGAACAACTGCAATGCAAGAT	11163
QY	6003	GTAACCAACCGCGCGACCAATATATGAGAAACACATTAAGTGCAGCCAGAATTCGACG	6062
Db	11164	GCAACCAACCAAGCTGATTCATATCTGAGACACCCCGCACAAAATGACCAAGATTTCCAG	11223
QY	6063	CGCGAGTTTCAAAACATCTTGGAACTGGCTGCTTGCACGTGTTTGGGGAGCATATCC	6122
Db	11224	CGCCCATCTCAAAAATCTTACATGAGATGGCTGTGTCCTTTTGGCGGGCGCTCGTCCG	11283
QY	6123	TCAATTGTGTAGACATTTATAGTGTGTGTGACGCTTATGCTTATMAACACAGTAAT	6182
Db	11284	TATTAATTAATAGACATTAATGATTTTGTCTTGACGATGATGTGACTGACACACAGAAAT	11343
QY	6183	GACTGAGCGCGGACACTGA-CATAGCGGTAAACTGATGTACTTCCGAGAAAGCTGGT	6241
Db	11344	GACCGGTAGCGCCCATATGACCGGACAGGAAAACCTGATGTACTTCCGAGAACTGATGT	11403
QY	6242	GCAATATGCCAGCGCGCGCT-----TGACACT	6268
Db	11404	GCAATATGATCAGCGCTGGTATATTAAGATCCCGGCTTACCGGCGCAATATGACAAACC	11463
QY	6269	AAAACCTGATGTAATTTCCGAGGAAGCACAGTGCATTAATCTGTGAGAGT	6318
Db	11464	AAAACCTGACGTAATTTCCGAGGAAGCGCAGTGCATTAATCTCGCAGAGT	11513

XX	AAV33476	standard; cDNA; 11703 BP.
XX	AAV33476;	
XX	10-DEC-1998	(first entry)
XX	Sindbis virus clone TR33	cDNA sequence.
XX	Bone marrow cell; alphavirus; Sindbis virus; cytokine; capsid;	
KM	reverse transcriptase-PCR; nsP1; nsP2; nsP3; nsP4; E3; E2; 6K; E1;	
KM	growth hormone; growth factor; interleukin; chemokine; enzyme;	
KM	ribozyme; antisense oligonucleotide; ss.	
XX	Sindbis virus.	
OS		
XX	Key	Location/Qualifiers
FH	5'UTR	1..59
FT		/*tag= a
FT	CDS	60..7601
FT		/*tag= b
FT		/product= "Non-structural polypeptide"
FT		/note= "Internal stop codon present at nucleotides
FT		5748-5750"
FT	CDS	7647..11384
FT		/*tag= c
FT		/product= "Structural polypeptide"
FT	3'UTR	11382..11703
FT		/*tag= d
XX	MO9836779-A2.	
XX		
XX	27-AUG-1998.	
XX		
XX	18-FEB-1998;	98MO-US02945.
XX		
XX	19-FEB-1997;	97US-0801263.
XX		
XX		
XX	(UTNC-) UNIV NORTH CAROLINA.	

XX Davis NL, Johnston RE, Simpson DA;
PI WPI: 1998-495361/42.
XX P-PSD; AAW70465, AAW70480, AAW70481, AAW70482, AAW70483,
DR AAW70484, AAW70485, AAW70486.
XX Expression of heterologous RNA in bone marrow cells - using a
PT recombinant alphavirus comprising a promoter operable in bone marrow
PT cells operably associated with heterologous RNA.
XX Example 6; Fig 6A-6B; 68bp; English.
XX The invention provides a method for introducing and expressing
CC heterologous RNA in bone marrow cells using alphavirus vectors.
CC The present sequence represents the Sindbis virus clone TR319 cDNA
CC sequence. The Sindbis virus cDNA encodes nonstructural and structural
CC polypeptides. The nonstructural polypeptide is, presumably,
CC post-transcriptionally modified into four different protein products,
CC namely nsP1 (AAW70464), nsP2 (AAW70465), nsP3 (AAW70480), nsP4
CC (AAW70481) proteins. The structural polypeptide is, presumably, post-
CC translationally modified into five different protein products, namely
CC capsid (AAW70482), E3 (AAW70483), E2 (AAW70484), 6K (AAW70485), E1
CC (AAW70486) proteins. The Sindbis virus cDNA was used in the method of
CC the invention. The inventors claim the transformed bone marrow cells
CC are useful for expressing a protein or peptide suitable for protecting
CC the subject against a disease such as a microbial, bacterial, protozoal,
CC parasitic or viral disease. The transformed bone marrow cells are also
CC claimed to be useful for expressing proteins and peptides such as
CC hormones, growth hormones, growth factors, interleukins, cytokines,
CC chemokines, enzymes, ribozymes or antisense oligonucleotides.
XX
SQ Sequence 11703 BP; 3306 A; 3051 C; 2913 G; 2433 T; 0 other;
Query Match 36.8%; Score 2388.4; DB 19; Length 11703;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3698; Conservative 0; Mismatches 1981; Indels 41; Gaps 8;
QY 580 ACTGACGGTATGAACCGGACCGCTATTTCTCATCGGAAACGGCCAAAGTCACTTC 639
DB 5746 ACTACTCAACCGGGAGTGGTATCATTTTTCAGCGACACAGGCCCTGGGCACTTGC 5805
QY 640 AACGAAATCAGTACGTCAATGTAACTACAAGAACTATTTGGATCGGCCCTGCATG 699
DB 5806 AAAAGAGTCCGCTCTGAGAACCAAGCTTACAGAACCACTTGGAGGCAATGCTCG 5865
QY 700 AGAAGTATACGCCCCCGCTCGATCTCGAAGAGAAATGTTACAGAAAGAACTGC 759
DB 5866 AAAAGATTCATGCCCCGGCTCGACAGTGAAGAGGAACTCAAACTCAGTATCC 5925
QY 760 AATATGCGCTCTGAGGAATAGAAAGAGTATCATCATCGAAAGTGAATATGA 819
DB 5926 AGATGATGCCACCGAAGCCAAAGTAGTACAGTCTCGAAAGTAAATAATCGA 5985
QY 820 AAGCAATTAACGCGAGGAGCTCATTTCTGGATTGGGACATATCTATCATCAGAAGTGA 879
DB 5986 AAGCATTAACCACTGAGGAGTACTGTCAGACTACAGACTATTAATCTGCGCAGCAG--- 6042
QY 880 ATCTGTGAGTGTAAAGTCAATTAATCTGTATCAACATCTACTCGTCAACGGTAAATTA 939
DB 6043 ATCAGCAGAAATGCTATTAAGTACCTATCGAAACATCTGTACTCCAGTGGCTA 6102
QY 940 ACAGGTTACATCTGACAGAGTGGCGGTAAACGTGCACTTATGTTATCCAAAGAAATT 999
DB 6103 CGAATCTCTCGAATCCAGTTCGCTGTAGCTGTGTAAACAATCTCTGATGATGAAC 6162
QY 1000 ACCCTACAGTACGAGTATGTATTAACAGATGAATAGATCGGATCTTACATGAGTGG 1059
DB 6163 ATCCGACAGTACATCTTATCAGATTACTGACGATGATGATCTTATGATATGGTAG 6222
QY 1060 ACGGCGCATCGTGTCTGTAGATACAGGCACTTTTGTCCGGCTAAACTGGAAGTACC 1119
DB 6223 ACGGACAGTGGCTGCTGATCTGCAACCTTCTGCGCCGCTTAAGATTAAAGTACC 6282

QY 1120 CAAAGAGCATAGCTATTTGCGAGCAGAGATAGATCAGCGCTCCATCGCTATACAGA 1179
DB 6283 CGAAAAAACAATGAGTATAGAGCCCGCAATATCCGAGTGGCTTCCATCAGCGATGAGA 6342
QY 1180 ATACATTACAAAATGATTTGGCTGCGAGCTACTAAAAGAAATTCGAACCTCAATGTC 1239
DB 6343 ACACGCTACAAAATGCTCATTTGCGGCACTAAAGAAATTCGAACGTCACGACAGATGC 6402
QY 1240 GAGATTAATCTGCTTATGATTTCCGCGGCATTTTAATGTTATGTTTCAAGAAATACCAT 1299
DB 6403 GTGAACGCGCAACACTGAGACTCAGCAGCATTAATGCGAATCTTTCGAAATATGCAT 6462
QY 1300 GCATGATGAGTCTGGGATACCTTTGGCATTAACCTATTCGCTACTACAGAGAAGC 1359
DB 6463 GTATATGCGATTTGGAGAGAGTTCCTCGGAAGCCATTAAGATTAACACTGAGTTTG 6522
QY 1360 TTACGAATATGTGACAAAGCTGAAAAGGCGAAAGCAGCAGCATTTGCGAATATCTC 1419
DB 6523 TCACCGCATATGTAGCTAGACTGAAGAGCCCTAAGGCCCGGACATTTTGGCAAGACGT 6582
QY 1420 ATATATCAAAACGTTGCGAGAGATACATGATGATCAATTCGTCTATGATCTTAAGAGAG 1479
DB 6583 ATATATTTGATCCATTTGCAAGAGTGCCTATGATGATTTGCTATGAGATGAAAGAG 6642
QY 1480 ATGTCAAGTATCTCCCGGACGAAACATACAGAGAGCGGCTAAGTGCAGGTTATTC 1539
DB 6643 ACCTGAAGATTAACACGAGCAGCAACGAAACACAGAAAGAACCGAAAGTACAGTATAC 6702
QY 1540 AGGCTGAGATCCCTTGTCTACCGCTTACCTTTGCGGAGTCCATCGGAAATGATCCGTA 1599
DB 6703 AAGCCGAGAAACCCCTGCGACATGCTTATATGCGGAGTTACCGGGAATATGATGCTA 6762
QY 1600 GACTGAATGCGGTCTTCTGCGCAATATCTCTTTCGACATGTACGCGGAAGATT 1659
DB 6763 GGGCTTACGCGCTTCTGCTTCCAAACATTCACAGCTTTTGAATGCTGCGCGAGATT 6822
QY 1660 TTGATGATTAATTTGCTGAACATTTCCACACGCGGACCCAGTATGGAAGAGCAATCG 1719
DB 6823 TTGATGATTAATTAATGCAAGAACTTTCAGAGAGGCAACCGGATCTGAGACGATATCG 6882
QY 1720 CGTGTGTTGTAAGGAGAGAGCAGCGCTATTCGCACTTTGGGCGTTGATGATCTTGAAG 1779
DB 6883 CATCATTTGCAAAAGCCAGAGCAGCGCTATGCGTTAACCGGTGATGATCTTGAAG 6942
QY 1780 ACTTAGGTGACCAACCGCTCTTGAATTTGATAGAGGCGGCTTCCGCAATATCAT 1839
DB 6943 ACCTGGGTGATCAACCACTACTGATTCAGTGCAGTGGCTTTGGAGAAATATCAT 7002
QY 1840 CTGTGACCTACCTACAGAGAGAGGTTAAATTTGGTGCATGATGAATCCGATATG 1899
DB 7003 CCACCATCTACTACGAGGATCTGTTTAAATTTGGGAGGATGATTAATCCGAAATGT 7062
QY 1900 TCTTAACGCTGTTTGTACACACACTAGTCAATATCATATGATTTGTCAGAGTACGTTG 1959
DB 7063 TCTTCAACATTTTGTGTAAACAGAGTTTGAATGCGTTATCGCAGAGAGTACTAAG 7122
QY 1960 AAGGTTAACAGTCAAGCTGCGGCGCTCTATTCGCGACAGTAACTATGTCATGCTG 2019
DB 7123 AAGGCTTAAAGAGTCAAGTGTGCAAGCTTCAATTCGCGACGACAACTATCATAGAG 7182
QY 2020 TCGTCTCCGACCTTGAATGGGAGAGATGGGCACTTGGCTGAACATGGAAGTAAAAA 2079
DB 7183 TAGATATGACAAAGAAATGCTGAGAGGTCGCGCACCTGCTTCACATGAGATTAAGA 7242
QY 2080 TTATGATGAGTATTTGATGATCAAGCACTTACTTGTGTGGAGATTTATCTGTGTG 2139
DB 7243 TCATGACGAGATCATGAGTGTGAGAGACCACTTACTTCTGCGGAGATTTATCTTGAAG 7202
QY 2140 ACCAGATTAACGAGCAGCTGCAAGTGCAGACCTCTTAAAGAGCTTTTAAAGTTG 2199
DB 7203 ATTCTGTTACTTCCACAGCGCTGCGGCGGAGACCCCTGAAAGAGCTTTTAAAGTTG 7262

OY	2200	AAAAACCATTCGCACTGCATGATACCCNAAGCTGGAACCGCCGCGGCACTGCATGATG	2235
Db	7363	GTAACCCGCTCCCAACCCGACGAGCAGAGAGAAAGAAAGAGAGCGCCCTCTGCTAGATG	7422
OY	2260	AAGCAATCGATGGAACAGAAATTGGAATTACGAGCGAGTTAGTGAAGGCGGTGAATCA	2319
Db	7423	AAACAAAGCGGTGTTTAGAGTAGTATTAACAGGCACTTTAGCACTGGCCGTGACAGACC	7482
OY	2320	GATACGAGATCACTACGCGAGGCGCTGATCATCAGCTCTCTGTCCAGCTTAGCCGNAACG	2379
Db	7483	GGTATGAGTAGACAATATTACCTGTCTCTACTGGCAATTGGAACCTTTTGCCAGAGCA	7542
OY	2380	TTAAGAACTTCAAGAGCATTAAGAGGAGACCCCAATACCTCTACGCGTGACCTAATATG	2439
Db	7543	AAAGAGCATTTCCAMCCCACTCAGAGGGGAAATTAAAGCATCTCTACCGTGTCTTAATAGT	7602
OY	2440	TGACGTATGACAGCACTAC-----CCACGGGCGAAGTGTTCATATCC	2486
Db	7603	CAGCATAGTACATTTTCATCTGACTAATACTACACACACACCACTGAAATGAGGATTTCT	7662
OY	2487	CTCAGCTGAACCTTTCACACAGTTTACCCCTACAAATCCGATGGCTTACCGAGATCCAAC	2546
Db	7663	TTAATCATCTCGGCGCGCGCCCTTCCGGGCCCACTGCAATGAGAGCGCGGAGAA	7722
OY	2547	CTCTAGGCGCGCGTGAAGCCGTTTGGCCGCCCGCTGGCTGCTAAATCGAAGTCTTA	2606
Db	7723	GGAAGCAGCGCGGCCGATGCTGCTGCCCAACGAGGCTGGCTTCTCAAAATCCAGCACTGA	7782
OY	2607	GGAGGTGATATAGTCAACTTGAACCTTCANCAACAGTCACTAATCCGCGCGCAGGTCCA-	2665
Db	7783	CCACAGCGCTCAGTCCCTCCTAGTCAATGGAACAGGCACTAAGACTTAACTCCACGCTCAC	7842
OY	2666	-----CCGCCAAAGAAAGAAAGAGTGTGCTCTAAGCCAAAACCTTCTCAGCCTTAAAGA	2702
Db	7843	GCCGCGCACCGGCGCCAGAAAGAGAGGCGCCCAACCAACACAGCAAGCCGAAGAACCA	7902
OY	2721	AGAMCGACCAACCCAAAGAGCAAAAC--GCAGCTTAAACAGGGAACGACACGTA	2777
Db	7903	AAACCGCAGAGAAAGAAAGAAAGCAACTGCAAAAACCCGAAAGAGACAGCGCA	7962
OY	2778	TGTGTATGAAGTTGGAGTCCGACAAAGACATTTCCGATCATG---CTGAACGGCCAACTGA	2834
Db	7963	TGGCACTTAAAGTTGAGGCGCCGACAAATGTTTGCAGTCAAAAGCAGGACGGAATGTCA	8022
OY	2835	ATGATATGCTCGCTTGTCCGAGGAGAGGCTGATGAACCACTCCGACCTTGAAGGAATA	2894
Db	8023	TCGGGCAAGCACTGGCCCATGGAAGGAAGTATATAACCTCTGCACTGTGAAGGAACCA	8082
OY	2895	TTGATTAATGACCAATTAGCGGCGCTGAATTGAAGAAAGCTAGCATGTACGACTTGAGT	2954
Db	8083	TCGACCAACCTGTGTGCTATCAAAAGCTCAAAATTTACAAATCGTCAAGCATTAAGCATGAGT	8142
OY	2955	AACGGGAGCTTCCCGAAGACATGAATACAGACAGCTGCAGTACACCGGCAAAACCAAC	3014
Db	8143	TCGACCACTTGCACACTCAACATGAAGAAATGAGGCAATTAACCTTACCACTGAAACACCCG	8202
OY	3015	CGGCGTTTCAACACTGCGACCAACCGCGGAGTCCAGTATGAGAAATGGAGATTTACCGTAC	3074
Db	8203	AAGATTTCTTAATCTGSCACCAACGAGCGGTGCACTATAGTGAAGTGAATTTACATCC	8262
OY	3075	CGAGGAGAGTGGCGGGAAGCGCACAGCGAAGACCGATCTGTGACAAACAGAGGACAG	3134
Db	8263	CTCGGCGAGTGAAGAGGACAGAGGACACGCGTCTCGATATGATTAATCTCCGGTGGG	8322
OY	3135	TTGTGCGTATGTCTTAAGAGGTGCAATTGAAGGACCGCTTACGGCGCTTTCAGTGGTCA	3194
Db	8323	TTGTGCGAATGTCTCTGCTGAGCTGTATGAAGAAACGAATCGCCCTTTCGTCGTCA	8382
OY	3195	CTTGGAAACAGAAAGGGGTGACCATTTAGGATATACCCCGAAGGTTCTGAACCGTGGT---	3251
Db	8383	CCTGGAAATGTTAAAGGAAGACAAATTAAAGACGACCCCGGAAGGACAGAAAGATGTGCG	8442
OY	3252	-----CACTAGTTACAGCGCTATGCGTCTTGCATGTGCACTGTTCCATGCGCAAAAC	3305

Db	8443	CAGCACCACTGTGTACGGCAATGTGTTGCTCGGAAATGTGAGCTTCCCAAGCCACGGCC	85020
Qy	3306	CACCCGTGTGTATTACTGACGGCCAGAACGAACTCTGACGTCTCGAAGAGAACTGCG	3365
Db	8503	CGCCCAKATGTCTTAATCCGGCGAACCTTCCAGAGCCCTCGACATCTTTGAAAGAACTGA	8562
Qy	3366	ACAAATCCAAATTACGACACGCTGCTGGAAACGTTTGAAATGTCATAC--GCCGC	3422
Db	8563	ACCAATGAGGCTTAAGATACCTGCTCAATGACATATGGCGGTGGATTCGCTGGCAGAA	8622
Qy	3423	CCAAACGAAGATTACCGATACCTTCACTGACCAATGCTTACTGTGGGTTCTGCCCT	3482
Db	8623	GCAAAAGAAAGGTCACTGACGACTTTCACCTGACACACCCCTTCTTGGGCAATGCTCGT	8682
Qy	3483	ATTGCAAGCACTCAACGGCGTGTTCAGGCCCAATPAAATGAGAACGTGTGGGACGAAT	3542
Db	8683	ACTGCCACCACTACTGAAACCGTGTCTTACGCCCTGTTAAAGATGACAGAGTCTGGAGCGAG	8742
Qy	3543	CTGATGATGATCGATTAGAAATCCAGGCTCTGGGCAATTGCGCTAGCAATCAGGACGCA	3602
Db	8743	CGGACGATTAACCAATAGCATACAGACTTCGCGCCAGTTTGATTAAGACCAAGCGGAG	8802
Qy	3603	CTGCGGATGTCAACCAATTCGCTTACATGTCTTTCGACGACGACCAATGACATCAAGAG	3662
Db	8803	CAGCAAGCGCAAAAGTACCGCTCAATGTGCTTGAGACAGGATCACACCGTTTAAAGAG	8862
Qy	3663	ACAGTATGGAGAAATAGCTATCAGCAATCTGACCCCTGCGCTGTTGGCCACAAG	3722
Db	8863	GCACCATGATGACATCAAGATTAGACCTCTAGGACCGGTGTAGAGGCTTAGCTAACAG	8922
Qy	3723	GGTACTTCTCTGTAGCTCAATGTCTCTCAGGTGACAGTGTAAACGTCAGTATCAGGACG	3782
Db	8923	GATCTTCTCTCTGGCAAAATGGCCCTCAGGGGACAGGTTAAGGTTAGCATAGTAGTA	8982
Qy	3783	GAGCATCTGAGAAATTCATGACCGGTGAGAAAAAGATCAGAGAAAGTTGTCCGTAGAG	3842
Db	8983	GCAACTCAGCAACGTCACTGTACACTGGCCGCAAGATPAAACCAAAATCGTGGGAGCGG	9042
Qy	3843	AGGAGTCTTGTCCACCCGCTGCATGGAAGCTGTAAAGTGCAGTTTACGATCACT	3902
Db	9043	AAAAATATGATCTACTCTCCGTTCCAGGTAAATTAATCTTGTGACAGTGTACGACGCTC	9102
Qy	3903	TGAAGAGAGCTCTGCGCGGTACATTAACATGTGACAGGCGACGCCACGCTGTAAAT	3962
Db	9103	TGAAAGAAACAACTGCAAGGCTAATCACTATGACAGGCGGGACCGCAACGCTTATACAT	9162
Qy	3963	CTTATCTGAGAAAGCGTCAGGCGCAAGTGTACATTTAAACCACTTCTGGCAGAACTCA	4022
Db	9163	CTTACTGTGAAGATTCATCAGGGAAGATTTTACGAAAGCGGCATCTGGGAGAACTTAA	9222
Qy	4023	CCTAAGAAATGTAATGTGTGGGCACTACAGCAACAGTATCTGTAGAGACCGCGAAACGAATGA	4082
Db	9223	CGTATGATGTAAGATGTGGGCACTACAGCAACGGAACCGTTTGCACCCGCAACGAAATCA	9282
Qy	4083	ACGCGTCACTTAAAGCAAAACAGTGCATTGCTCTCAAGAGCGACCAACGAAATGGGCTCT	4142
Db	9283	CTGCTTGCACCGCATCAAGCAAGTGTGCTGTCTTAAAGAGGACCAACGAATGGGCTCT	9342
Qy	4143	TCAACTGCGCGGATCTTATTAAGGCAACAAGCACTCAATGCAAGGTAAATTCGACATTC	4202
Db	9343	TCAACTCACCGGACTTGTATCAGATACGACGACCAAGCGCCCAAGGGAATTTGATTTGC	9402
Qy	4203	CATTCCGCTTGACACCGACAGTCTGCGCGGTTCCGTTAGCTCAACCCCTTACAGTCAAGA	4262
Db	9403	CTTTCAATGATATCCAGATACCTGTAGTGTCTCTGTGCTCCACGCGCCGAAATTAATAC	9462
Qy	4263	AGTGTCTCAAAGGATCACCTTCACTGTGACTGTGACATGACCAACATTTGCTGCAACGA	4322
Db	9463	ATGCTTTTAAACATATGAGCTTCAATTAGATACAGACCACTTGACATTTGCTACCAACA	9522
Qy	4323	GAAATATGGGCTGCGAGCAGACGCAACAGCAGATGATTAACGGTCTTACATCCAGGA	4382

Dh 9523 GGAGACTAGGGGCAAAACCGGAAACCAACTGATGATGTCGGAAGACGCTGCAAA 9582
Qy 4383 ATTTTCTGTTGGGGGCGAAGGGGTGAGTACGATGAGGGGTAAACCATGAACCACTGAG 4442
Db 9583 ACTTCACCTGACCGAGATGCTGGAAATACATATGAGGAATATGAGCCACTGTGGG 9642
Qy 4443 TCTGGGCGCAGAGTCCGACCAAGCGACCCACATGATGCGCGCATGATCATCTCC 4502
Db 9643 TCTATGCCCAAGAGTCAAGACCAAGAGACCTCAAGATGCGCACAGAAATATATACAGC 9702
Qy 4503 ACTATTTATCATGCGCATTCAGTCTACACTGTCTATGCTGTGTGTGTGTGTGTGT 4562
Db 9703 ATTAATCACTATGCGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9762
Qy 4563 TCCGTGAGGACCTGATCATCTGACAGCTTGCATGCGCAAGCAAGAGACTGCGTGA 4622
Db 9763 TGATGATTTGGGTGTAACGTTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9822
Qy 4623 CGCCATACGCGCTTGCAACCGAACCGAACCGTACCCACAGCATTTAGCGGTTTGTCTGA 4682
Db 9823 CGCCATAGCGCTTGGCCCGCAAGCGCGTAATCCCACTGCGCTGCACTCTGTGTGTGTG 9882
Qy 4683 TTGCGCCAAACCAAGCTGAAACATTTGAGAACTTTGAACTTGTGTGTGTGTGTGTGT 4742
Db 9883 TTAGGTGCGCAATGCTGAAACGTTTCAACGACCATGATGATCTTGTGTGTGTGTGTGT 9942
Qy 4743 AACGTTTCTGTGGGCAAGTGTGATCTCTGTGGAGCGCTTGTATTTCTGTGTGTGT 4802
Db 9943 AGCGTTTCTGTGGGTCAGTGTGATCTTGTGGCGCGCTTGTATTTGTGTGTGTGTGTGT 10002
Qy 4803 GCTTTTATGCTGATGCTTTTATTTATTTGTGTGAGCGTGTGTGTGTGTGTGTGTGTGT 4862
Db 10003 GCTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10062
Qy 4863 CTTTGAACATGCGACCACTGTGCAAAATGTTCCGGGGATCCGTAATAGCGCTGTGTGT 4922
Db 10063 CTTTGAACATGCGACCACTGTGCAAAATGTTCCGGGGATCCGTAATAGCGCTGTGTGT 10122
Qy 4923 AACGCGAGGTATGGCGCACTTAACCTGAGATCAAGGTGTGTGTGTGTGTGTGTGTGT 4982
Db 10123 AAAAGGAGGGGTATGCGCTCAATTTGAGATCACTGTATGCTCTCGGAGGTTTGTGC 10182
Qy 4983 CTTTCAATTAACAGAGTACCTGACCTGCAAAATTTCAACAGTATCTTTCACACAG 5042
Db 10183 CTTTCAATTAACAGAGTACCTGACCTGCAAAATTTCAACAGTATCTTTCACACAG 10242
Qy 5043 TTAATGCTGTGGGTCCTCGAGTGAAGGCACTCTCAAGCGGATTAACATGCGCG 5102
Db 10243 TCAATGCTGTGGGTCCTCGAGTGAAGGCACTCTCAAGCGGATTAACATGCGCG 10302
Qy 5103 TTTTGGCGGTGTATCCCTTCAATGTTGGGAGGCGCAATGCTGTGTGTGTGTGTGTGTGT 5162
Db 10303 TCTTGGAGGGGTATCCCTTCAATGTTGGGAGGCGCAATGCTGTGTGTGTGTGTGTGTGT 10362
Qy 5163 ACACACAACTGAGTGAAGCGTACGTCGAGTTCGCTCGACAGTCACTATAGATCAGAG 5222
Db 10363 ACACACAACTGAGTGAAGCGTACGTCGAGTTCGCTCGACAGTCACTATAGATCAGAG 10422
Qy 5223 TCGCACTAAAGTTCAACAGCTGCTGAAAGTGGCGCTGTAAGTATATGCGCAACA 5282
Db 10423 AGGCGATTAAGGTCACTGCGGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 10482
Qy 5283 CCACCGCGCACCTGATACGTTTGTCAATGCGGTCAAGCGGATTTCCACAGGGACCTGGA 5342
Db 10483 CTACCACTTTCTTGAAGTGTGATGCGAAAGGAGTCAACCGAAGCGTAAAGACTTGA 10542
Qy 5343 AGGTATAGCAGAGCGCATATCAGCGCTTTTTCACCTTTGACCATTAAGTCTGTATCA 5402
Db 10543 AAGTATAGTCTGACCAATTTCAAGATGTTTACGCACTTGTGATTAAGTCTGTATCC 10602
Qy 5403 GAAAGGGGCTTTTATCAACTACGATCTTCTGTGATGAGCTATGAAACAGAGCGT 5462
Db 10603 ATCCGCGCTGTGTATCAACTATGATCTTCCGGAATATGAGGATGAAACAGAGCGT 10662

Qy 5463 TCGCGATATTTCAGACATCTCGCTTGATGCTACAGACATAGAGCCCGCATGACATAC 5522
Db 10663 TTGAGACATTTCAAGTACTCTCTTGACTAGCAAGATCTCATCGCAGACAGACATTA 10722
Qy 5523 GCGTCTGAAGCTTCTGTCAAGAACATCCAGTCCCTCTACACCAAGCATATCAAGGT 5582
Db 10723 GCGTCTGAAGCTTCTGTCAAGAACATCCAGTCCCTCTACACCAAGCATATCAAGGT 10782
Qy 5583 ATGAATGTGGAAGAACATCCAGTCCCTCTACACCAAGCATATCAAGGT 5642
Db 10783 ATGAATGTGGAAGAACATCCAGTCCCTCTACACCAAGCATATCAAGGT 10842
Qy 5643 AAATGAGTGAAGCTTCTGTCAAGAACATCCAGTCCCTCTACACCAAGCATATCAAGGT 5702
Db 10843 AGATGAGTGAAGCTTCTGTCAAGAACATCCAGTCCCTCTACACCAAGCATATCAAGGT 10902
Qy 5703 TTGACATCTCTGATGACGCTTTTGTGATCATCAAGAACATCCAGTCCCTCTACAGTGA 5762
Db 10903 TTGACATCTCTGATGACGCTTTTGTGATCATCAAGAACATCCAGTCCCTCTACAGTGA 10962
Qy 5763 GCTGCACTGAGACGATGATTTATTTCTGACAGCTTGTGTGTGTGTGTGTGTGTGTGT 5822
Db 10963 AATGTGAGTCAAGTGTGATCTTATTCAGACAGCTTGTGTGTGTGTGTGTGTGTGTGT 11022
Qy 5823 ACAAGCTGACAGGAGGACATTTGTCAAGTGTCACTCCCATCCACAGACAGTGTGTGTGA 5882
Db 11023 ATGTATTCGACCGCGAGGATCAATGCGCTTACATTTGTGACAGACAGCACTCTCC 11082
Qy 5883 AGAAGCGACACACATGATGATCCGTGAGGACAGATTAACATTTTATGACATCGA 5942
Db 11083 AAGAGTGCACAGTATGATCTCTGAGAAAGAGCGGTGACAGTATCACTTTAGCACGCGGA 11142
Qy 5943 GCCCAAGCAAAATTTATTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6002
Db 11143 GTCCACAGGGAATTTATTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11202
Qy 6003 GTAAACCAAGCGGACGACATATTAATGAGAACCACTAAAGTGTGACCAAGAAATTCAGG 6062
Db 11203 GTAAACCAAGCGGACGACATATTAATGAGAACCACTAAAGTGTGACCAAGAAATTCAGG 11262
Qy 6063 CGGCACTTTCAAACATCTTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6122
Db 11263 CGGCACTTTCAAACATCTTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11322
Qy 6123 TCAATGTTGAGACTTATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6182
Db 11323 TCAATGTTGAGACTTATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11382
Qy 6183 GACTGACCGGACACTGA-CATAGCGGTAAACTGATGATCTTCCGAGGAAGCGTGT 6241
Db 11383 GACTGACCGGACACTGA-CATAGCGGTAAACTGATGATCTTCCGAGGAAGCGTGT 11442
Qy 6242 GCATTAATGACAGCGCGCT 6261
Db 11443 GCATTAATGACAGCGCTGT 11462

RESULT 5
AAV3474
ID AAV3474 standard; cDNA; 11717 BP.
XX
AC AAV3474;
DT 10-DEC-1998 (first entry)
XX
DE Girdwood S.A.virus cDNA sequence;
XX
KW Bone marrow cell; alphavirus; Girdwood S.A.virus; cytokine; capsid;
KW reverse transcriptase-PCR; nsP1; nsP2; nsP3; nsP4; E3; E2; E1;
KW growth hormone; growth factor; interleukin; chemokine; enzyme;
KW ribozyme; antisense oligonucleotide; ss.
XX

Db	6558	ACCTGGGGTGTGGACCAACCACTCACTCCACTTGGATCGAGTGGCGCTTTGGAGAAATATCAT	7017
Qy	1840	CTGTGGCACTTACCTTACAGAAAGAGGTTTAAATTTGGTCAGATGAAATCCGGTATGT	1899
Db	7018	CCACCCCATCTGCCACGGGTACCGGTTTCAATTCGGGGGCGATGATGAAATCCGGAAATGT	7077
Qy	1900	TCTTAAAGCTGTTTGTGCAACAACCTAGTCATATTCATATGATTGTGACGAACTTACAGT	1955
Db	7078	TCTCCACGCTCTTTGTCAACAACAGTTCTGAATGTCTGTTATGTCCACGACGAGATTTTGAAG	7137
Qy	1960	AACGGTTTAAACAGCTCAGCGTGGCGGCGCTCTATCGGCGAGATTAACCTATGTGATG	2019
Db	7138	AGCGGCTTAAACGTCCTAAATGTGACAGATTATTCGGGAGACGACACTCATACAGSAG	7197
Qy	2020	TCTGTCCTCCAGACCTTGATGGCGAGAGATGCGCCACTTGGCTGAACTGSAAGTAA	2079
Db	7198	TAGTATCTGACAAAGAAATGGCTGAGATGTGACACTGGCTCAACATGAGGTTTAA	7257
Qy	2080	TTATTTGATGCACTTATTTGGTATCAAAGAACCTTACTTCTGTGGGGGATTTATCTGTGTG	2139
Db	7258	TCATTGAGGCAAGTATCGGCGAGAGACCGCTTACTTCTGTGGGTGGATTCATCTTGCAG	7317
Qy	2140	ACCAAGATTAACGAGCGACGCTGCGAGAGTCCGACAGCCCTTAAAAAAGCTTTTAAAGCT	2199
Db	7318	ATTGGGTATCTCTCCACAGCGTGTGCGCTGGCGGACCTTGAAAAAGCTGTTTAAATTTGG	7377
Qy	2200	GAAAAACATTGCCAGTCGATGATACCCAGACTGCGACCGCGCGGCACTGATGATG	2259
Db	7378	GTAACACGCTCCACGCGCAGCAGCAGACGAAAGACAGAAAGACGCGCTCTGCTAATG	7437
Qy	2260	AAGCAATGCCATGGAACAGAAATTGGAAATTAACGAGAGATTTAGTAAGCCCTGAAATCCA	2319
Db	7438	AAACAAAGGCGTGTGTTTGAATAGGTAATACAGACACTTTCAGATGTCGCTGCAATCTC	7497
Qy	2320	GATTCGAGATCATCTGGCAGGCGGTATCATCGTCTCTGCAACGTTTACCGGAAAGCG	2379
Db	7498	GGTTTGAAGTGAACAACATCACTCTGTCTGCTGGCACTTAAAGACTTTTCCAGAGCA	7557
Qy	2380	TTAAGAACTTCAAGGACATAAGAGGAGGCCCAATCACCTCTAGCGCTGACCTTAATAG	2439
Db	7558	AAAGAGCACTTTCAGGCACTCAGAGGGGAAATTAAGCACTCTAGCGTGTCTTAATATGT	7617
Qy	2440	TGACGTAGTA-----GACACGCACTTACCAACCGGCGAATGTTTCCATACC	2486
Db	7618	CAGATATGCACTTTCATCTGACTATATCCACAACACCAACCAACATGATAGAGATTC	7677
Qy	2487	CTCAGCTGAACTTTCCACACAGTTTAAACCTAAGAAATCGATGCTTAACGAAATCCAAAC	2546
Db	7678	TTTAACATGCTGGCGCGCGCCCTTCCCGGCGCCCACTGCAATGTGAGGCGCGGAGAA	7737
Qy	2547	CTCTTAGGCGCGCTGAGAGCGGTTTCGAGCCCGCTGGCTGCTCAATTCAGATCTT	2606
Db	7738	GGAAGCAGGCGGCGCCCAATGCTCTGCGCCCAATGGGCTGTGCTTCCAAATTCAGCACTGA	7797
Qy	2607	GGAGGTGATATGCTCAACTTGAACCTTCAACA-----ACGATCACCTTAATCCGCGCGAG	2666
Db	7798	CCACAGCCGTGATGCTCCTAGTCTATTGACAGGCACTAGACCTTCAACCCCAAGCCAC	7857
Qy	2661	GTCCACCGCCCAAGAAAGAAAGAGTGTCTTAAGCCAAACCTATCAAGCTTAAAAAG	2720
Db	7858	GCCCCGCGCGCGCCAGAAAGAGCGGCGCCAAAGCAACACGAAAGCGAAGAAACCA	7917
Qy	2721	AGAAAGCAAGCCAAAGAGAGCGAAAC---GCAAGCTTAAACCAAGGAAAGAGCAACGTA	2777
Db	7918	AAACACAGGAGAAAGAAAGAAAGCACTTCGCAAAACCCGAGAAAGAGACAGCTA	7977
Qy	2778	TGTGTATGAAGTTGGAGTCGACAAAGACATTTCCGATC---ATGCTGAACGGCAAGTGA	2834
Db	7978	TGGCACTCAAGTTGGAGGCGCAGACAGACTGTTGCAGCTCAAAAATGAGGACGAGATGTCA	8037
Qy	2835	ATGATATATGCTCGCTTTCGAGAGAGCGTGAATGAACCACTCAAGTTAAGAAAA	2894
Db	8038	TCGGGACACGACTGGCCATGAGAGGAAAGTTATGAAACCACTCACTGTAAGAAAGAACTA	8097

QY	2895	TTGATATAGACATTTAGCGCCCGTGAATTTGAAGAAAGCTAGCATGTATCGACTTTGGAGT	2959
Db	8098	TTGACACACCCCTGTGCTATCAAAAGCTCAAAATTCACCAAGTCGTTCAGCATATCACATGAGT	8157
QY	2955	ACGGGCAAGTTCCTCCAGAACATATGAATTCAGACAGCTGCAGTATACCAAGCGACAAACAC	3014
Db	8158	TTCGACAGTTCCCGGTCAACATATGAAAGTAGAGGGTTTCACTTATACCAAGCGAACACCTG	8217
QY	3015	CGGGCTTCTACACTGCGTCAACGCGCGCAGTCCAGTATGAGATGGAGATTTACCGTAC	3074
Db	8218	AAGGTTTTCATACCTGGCACACCGAGCGGTGCGATATGAGGATTAATTCATCC	8277
QY	3075	CGAGAGAGTGGGGGGGAAAAGGCGACACGCGAACAACCGATCTTGGACACAGAGCGAG	3134
Db	8278	CCCGCGAGTATGAGAGCGAGAGAGACAGATGTCGTCCGATTTATGATATCTCAGGCCGGG	8337
QY	3135	TTGTGGCTATTGTTCTTAGAGAGTCAATGAGGCGACGCGTACGCGCTTTCACTGTGCA	3194
Db	8338	TTGTGCGCATATGTCCTCGAGGGGCTGATGAGGAAACAAAGATGCTCTTTCGTCGTCA	8397
QY	3195	CTTGGAACCAAGAAAGGGTGCATTTAGGATATACCCCGAAGGTTCTGAACCGGCT---	3251
Db	8398	CCTGAATATGCAAAAGGAAAGCAATCAAGACAAACCCCGAAAGGACAGAAAGTGGCTTG	8457
QY	3252	-----CACTAGTTACAGCGCTATGCGTCTTTGAAATGTCACTTCATCCATCGACAAC	3305
Db	8458	CAGACCACTGCTGACCGGCACTGTCGTTGTTGTTGAAAGTGAAGCTTCCATGCAATCGCC	8517
QY	3306	CACCCGTGTCTATTCCTACTGACGCGCAAGCAACATCTGCAGTCTCGAAGAAACGTCC	3365
Db	8518	CGCCCACTGCTTACACCGCGAACCATTCAGAGGCTTTGACATCTTTGAAGAAAGCTGA	8577
QY	3366	ACAATCCAAATTTACGACAGCTGCTGGAGAACGCTTGAAT---GTTCATCAGCGCGC	3422
Db	8578	ACCAAGAGGCTTAGACACCTTGCTCAACGCAATTTGCGGTGGGATCGTCGGCGAA	8637
QY	3423	CGAAACGAAGATTACCGAGTACCTTCACTGACAGTCCCTACCTGGGGTTCTGCCGT	3482
Db	8638	GCAAAAGAGGCTCACTGACGACCTTACCTTTCACACCGCGTACTTGGGCAACATGCTGT	8697
QY	3483	ATTGACAGCATCMAACGCGCGTTTCAGCCCAATTAATAATTGAAAGAGTGGGAGCAAT	3542
Db	8698	ACTGCACCAATACGAACGCTGCTTAGCCCGATTTAAGATCGACAGGCTGGGATGAG	8757
QY	3543	CTGATGATGATGCTATTGAATCCAGGCTTCGCGACATTTGGGCTACATTCAGGACGCA	3602
Db	8758	CGGACGACAAACCACTATCGATACAGATTCGCCGCCAGTTTGGATACGACCAAGCGGAG	8817
QY	3603	CTGGGAGTGTACCAAAATTCGGTTACATGTTTTCGACACAGACCATGACATCAAGAG	3662
Db	8818	CAGCAAGCTCAATATAGTACCGGCTACATGTCGTCGAGCAGGATCATACGTTCAAGAG	8877
QY	3663	ACAGTATGAGAAATATAGCTTACGACATCTGGAACCTCGCGCTGCTTTGGCCACAAG	3722
Db	8878	GCACTATGATGACATCAAGATCAGACCTGAGACCGGTGTAAGAAAGCTTAGCTACAAAG	8937
QY	3723	GGAATCTCTGTATAGCTCATATGTCCTTCAAGGTGACAGTGAACCGTACATACAGAGCG	3782
Db	8938	GATACTTCTCTCGCGGAAGTGTCTCAGGGGACAGGTAAACGTTTAGTATACCGAGTA	8997
QY	3783	GAGCATCTGAATAATTCATGACCGGTGAGAAAGATCAAGAGGAAGTTTGGTGGTAG	3842
Db	8998	GCAACTCAGCAAGCTCATGACAAATGGCCGCAAGATTAACAAAAATTCGTGGAGCGG	9057
QY	3843	AGGATACTTGTTCACACCCGCTCATGGAAGAGCTGTAAAGTCCAGTTTACATCACT	3902
Db	9058	AAAAATATGACCTTACCTCCGTTACCGTAAAGAAATTCCTTGACAGTGTACGACCTGC	9117
QY	3903	TGAAGGAGAGCTTCGCGGGTACATAACATGACAGGCCAGGCCACACGCGTATTAAGT	3962
Db	9118	TGAAGAAACACCGCGGCTACATCACTATGACAGAGCGGAGCGCACGCGCTATACGT	9177

QY 3963 CCTATCTGGAGAGCGTCAGCGGAAGTATACATTAAACCACTTCTGGCAAGACGTCA 4022
Db CCTATCTGGAGAGTATCATCGGAAAGTCTACGCGAAGCCACCATCCGAAAGAACATTA 9237
QY 4023 CCTAGGAATGTAAGTGTGCGACTACAGACAGATATCTGTAGCAGCGCAAGAAATGA 4082
Db CCTAGAGAGTGAAGTGTGCGACTACAGACAGATATCTGTAGCAGCGCAAGAAATGA 9297
QY 4083 ACGGCTGCACTAAGCAAAACAGTGCATTGCTACAAAGACCAAAAGAAAGGATCT 4142
Db CGGGCTGCAAGCGCCCATCAAGAGTGTGTGCTATTAAGGCAACCAAGAAAGTGTG 9357
QY 4143 TCAACTCGCGGATCTTATTAGGCAACAGACCACTCACTGCAAGGTAATTCACATTC 4202
Db TCAATTCGCGGATCTTATGATCAGACATGCGACCAACGCGCAAGGAAATTCATTC 9417
QY 4203 CATTCGCTTGACACCGACAGTGTGCGGCTTCCGTAGCTACACGCTTAAGTACGA 4262
Db CTTTCAAGCTGATCCGAGTACCTGTATGTCTGTTGCCACGCGCGCAAGTATAC 9477
QY 4263 AGTGTTCAAAGGATCACTCCACCTGACCTGACCTGACCAATGGACCAATTCGTGACAGA 4322
Db ACGGCTTAAACATCATCACTCTCTCAATTAGACACAGACCACTGACATTCGACCA 9537
QY 4323 GAAAAATGGGCTGCGAGCAGACGCAACAGAAATGATTAAGGCTTACATCCAGA 4382
Db GGAAGCTAGGGGCAATCCGGAACCACTAGATGATTCATCGAAAGAGGTTAGAA 9597
QY 4383 ATTTTCTGTGGGGGAGAAAGGCTGAGATAGTATGGGTTAACATGAACAGTACAG 4442
Db ACTTACCGTGCACGAGATGCGCTGGAATATATATGGGCAATACGAAACGCTAAGG 9657
QY 4443 TCTGGGCGCAGAGTCGCGACGAGGAGACCAATGATGAGCGGATGATCATCATC 4502
Db TCTATGCGCAAGATCTGCAAGAGACCTTCAAGATGCGCAACGAAATGATCAGC 9717
QY 4503 ACTATATCATGGGATCCAGTCTACATGTCATGTCGTGTGTGCTGCTGTGCTA 4562
Db ATTACTACATCGCCATCTGTGTACACCATCTTAGCGGTGACATCAGTGTGTGGA 9777
QY 4563 TCTGTGTAGGCACTGATATCAGACAGCTTGCAATGCGCAAGAGAGACCTGCTGA 4622
Db TGAATATGCGCTAACTGTGTGACATATATGCTGTAAAGCGCGCTGTGCTGCTGA 9837
QY 4623 GCGCATGCGGCTTGACCGCAAGCGATCCACAGACCTTAGCGGTTGTGCTGCA 4682
Db GCGCATATGCGCTTGCGCCCAATGCGGTATTCACATTCGCTGCGACTTGTGCTGTG 9897
QY 4683 TTGCGCAACCAACGCTGAAACATTTGAGAACTTTGAACCATCTGTGTTAAACAAC 4742
Db TTAGTTCGGCTAAATGCTGAACATTCACCGAAGCATGATGTTACTATGTGTGAACGCG 9957
QY 4743 AACGTTTCTGTGGCAACGTTGTGCAATCTGTGCAAGCGGTTGTATTCGTTCCGT 4802
Db AGCCATTCCTTGTGGTTCAGCTGTGTATCCCTGCGCGCTGTCAATGTTAATGCGCT 10017
QY 4803 GCTTTCAATGCTGCAATGCTTTTATTTATGTTGTGCAAGCGCTGTGCTGGGAAAGTGAAG 4862
Db GTTGTCTATGCTGCTGCTTTTATTTATGTTGTGCGCGCGCTTACCTGTGGAAAGTGAAG 10077
QY 4863 CTTTGAAATGTCAGCACTGTGCGCAATGTTCCGCGGATCCCGTATTAAGGCGTTGTG 4922
Db CTTAAGAACATGCACTGTGTTCCAAATGTGCGCAAGTACGTTAAGGCACTTGTG 10137
QY 4923 AAAGGCGAGTTTACGCGCACTTAACTGTGAGATCAAGGTGTCTCATGGAATTAACAC 4982
Db AAAGGCGAGGTTACCGCGCTCAATTTGGAATTAATGTCAATGTCGCGAGGTTTGC 10197
QY 4983 CTTCAACTAACAAGAGTACGATGCAATTTCCACACATCATCTTTTACACAGAG 5042
Db CTTTCAACCAACAAAGATACATCACTGCAATTAACACATGTGTGCTTCCCTTAAAG 10257
QY 5043 TTAATGTGCGGCTGCTGAGTGAAGGCAATCTCAAGGCGGATTAACATGCTCGCG 5102

Db 10258 TCAATGTGCGGCTGCTTGGAAATGTCAGCCGCGCTCAGCAGACTATACCTCAAG 10317
QY 5103 TTTTGGGCGGTGTGATCCCTTTCATGTGGGAGGCGCAATGCTTCTGTGACATGAGA 5162
Db TCTTGGAGGGGTGTACCTTTCATGTGGGAGGCGCAATGCTTCTGTGACATGAGA 10377
QY 5163 ACACAACTGAGTGGAGGCTGATGCTGAGTTCGCTCCAGACTGCACTATGATCAAGCAG 5222
Db ACAGCAATGATGAGGCGTATGCTGAAATTCATGACAGATTTGCCGATGCAAGCGCC 10437
QY 5223 TCGCACTAAAGTTCAACAGCTGCTGTAAAGTGGGCTGTGATTAAGGCAACA 5282
Db AGGCAATTAAGTGTGATCTGCGGATGAAGTATGAGTATGATGATGAGGAAACA 10497
QY 5283 CCAAGCGGCACTGATACGTTTGTGCAATGCGGTACAGCGGATTCCTACGGGACCTGA 5342
Db CTAACGATTTCTTAATATGTATGTAAGGAGTACACAGAAAGTCTTAAGGACCTGA 10557
QY 5343 AGGTATGACAGGCGGATATGAGCGGCTTTTTCACCTTTGACCATTAAGGTGCTATCA 5402
Db AAGTATGAGTGTGACCAATTTCAAGATGTTTACACATTTGATGACAAAGTGTATTC 10617
QY 5403 GAAAGGCGCTGTTTAACTAAGCACTTCCCTGATGAGCTATGAACCAAGAGCGT 5462
Db ATCGGCGCTGCTGTATCAACTATGATCTCCGGAATACGAGGATGAACCAAGAGCGT 10677
QY 5463 TCGGCAATTTCAAGATCTTCTGCTGATGCTACAGACATGATGAGCGGCACTATAC 5522
Db TTGAGACATTTCAAGCTCTCTTGTGATGACCAAGATCTCATCCGACAGACATTA 10737
QY 5523 GCGTGTAGGCTTCTGTCAAGAACATCCAGCTGCTCCCAACCCCAAGAGTATGAGGT 5582
Db GACTATGAGCTTCTGCGCAAGAGCTGATGCTTCCCTTACAGAGGCGGATTTGAT 10797
QY 5583 ATGAATGTGGAAGAACTCAGGACGACCCCTGCAAGAAACAGACACATTTGATGA 5642
Db TCGAGATGTGGAAGAACTCAGGCGCGCCCTGCAAGAAACAGGCTTCCGCTGCA 10857
QY 5643 AAATGAATGAGGCTTGTGAGGCTTAACTGTGCTTAAAGGAGACATCTTATCTGA 5702
Db AGATTTGAGTCAATTCGCTTGTGAGGCTGCTGCTTAAAGGAGACATTTCCATCTCTA 10917
QY 5703 TTGACATCCCTGATGACCTTTTGTGATGATGATGATGATGATGATGATGATGAT 5762
Db TCGACATCCCGACGCTGCTTATCAGAGATGACAGACATGCTGCTTCAACAGTCA 10977
QY 5763 GCTGCAAGTACAGACTGATTTATTTCTGAGACTTGTGTGTTCTTAACTTACAGT 5822
Db AATGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11037
QY 5823 ACAAAGTACAGAGGAGGACATGCTCAATTCCTCACTCCCATCCACAGAGCTGTTGA 5882
Db AATGTATCCGAGCGGAGAGAAATGCTGTATTCATTCGATTCGAGCAGAGAACCTTCC 11097
QY 5883 AGGAAGCAACCAACATGATGCTGCGGATGAGCAAGATTAACATTAATTGACATGA 5942
Db AAGATGTACAGTATGATGCTGAGAAAGAGCGGTACAGTACCTTACAGCAGCGGA 11157
QY 5943 GCCCAAGCAAAATTTTATGTTTCTATGCTGAGCAAGAGTCCAGTCAATGCTGAT 6002
Db GCCCAAGGCAATTTTATGTTTCTATGCTGAGTGTGATGATGATGATGATGATGAT 11217
QY 6003 GTAAACCAAGCGGCGACCAATTAATTTGAGAAACCAATTAAGTGAACCAAGATTTCAAG 6062
Db GCAAAACCAAGCTGACCATATGAGAACACCCCGCAAAATTAAGCAAGATTTCAAG 11277
QY 6063 CGGCAATTTCAAAACATCTTGAACATGCGGCTTGAACATGTTGGGGGAGCATATCC 6122
Db CCGCATTTCAAAACATCTTGAAGTGTGCTGCTGCTTTCGCGCGCGCTGCTGCTG 11337
QY 6123 TCAATGTTGAGACTTATGATGTTGCTGAGCTGATGCTTATTAACACAGCTAGAT 6182

Db 11338 TATTAAATATAGACTTAATGATTTTCTTGACAGCATGCTGACTAGCACAGCAAGAT 11397
Qy 6183 GACTAGAGGGGAGACAGTA-CATAGCGGTAAACCTGATCTTCGAGGAGACGGT 6241
Db 11398 GACCGCTACGCCCCATGACCCGACACGACAACTCGATGATCCGAGGAACTGATGT 11457
Qy 6242 GCATATGCCCAGCGCGCT-----TGACACT 6268
Db 11458 GCATATGCTACAGGCTGTATATTAAGATCCCGCTTACCGGGCAATATAGCAACACC 11517-
Qy 6269 AAAAAGTGTATTTTCCGAGGAGACAGTCAATATCTGTGCACTGT 6318
Db 11518 AAAAAGTGTATTTTCCGAGGAGACAGTCAATATGCTGCGCAGTGT 11567
RESULT 6
AAC64506 standard; DNA, 11703 BP.
AAC64506;
AC AAC64506;
XX 14-FEB-2001 (first entry)
DT 14-FEB-2001 (first entry)
XX Alphavirus Sindchiron virus nucleotide sequence.
XX Alphavirus; Sindchiron virus; SindchironLP virus; immune response;
KM infection; human dendritic cell; immunostimulatory; cytotoxic;
KM virucide; fungicide; antibacterial; antiparasitic; vaccine; cancer;
KM pathogen; antigen presenting cell; ds.
XX Alphavirus.
OS
XX W0200061772-A2.
PN 19-OCT-2000.
PD 14-APR-2000; 2000WO-US10722.
PF 14-APR-1999; 99US-0129498.
XX 09-AUG-1999; 99US-0148086.
PR 22-MAR-2000; 2000US-0191363.
XX (CHIR) CHIRON CORP.
PA
PI Polo JM, Dubensky TW, Frolov I, Gardner JP, Otten G, Barnett S;
PI Driver DA.
XX WPI; 2000-619231/59.
XX New alphavirus that infects human dendritic cells for use in generating
PT an immune response to pathogenic agents such as bacteria, viruses,
PT fungi, parasites and cancer and for biological assays -
XX
XX Claim 9; Fig 2B; 83p; English.
XX The present invention describes an isolated alphavirus (AV) which
XX infects human dendritic cells and is not of American Type Culture
XX Collection (ATCC) number VR-2526. AAC64506 and AAC64507 represent the
XX nucleotide sequence of the specifically claimed Sindchiron virus and
XX SindchironLP virus. The new AVs have immunostimulatory, cytotoxic,
XX virucide, fungicide, antibacterial and antiparasitic activities and
XX can be used in vaccines. The AVs are used to infect dendritic cells,
XX preferably human cells. A heterologous sequence can be introduced and
XX expressed in human macrophages or antigen presenting cells in vivo and
XX in vitro, for use in biological assays. The AV-based vector systems
XX are used to generate an immune response to cancer or a pathogenic
XX agent, such as, bacteria, fungi, parasites or viruses. The AV can be
XX used to infect human dendritic cells, macrophages or antigen
XX presenting cells that previously could not be infected using an AV or
XX AV variant. The AV vectors are targeted directly to antigen presenting
XX cells.
SQ Sequence 11703 BP; 3310 A; 3049 C; 2910 G; 2434 T; 0 other;

Query Match 36.8%; Score 2383.6; DB 21; Length 11703;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3695; Conservative 0; Mismatches 1984; Indels 41; Gaps 8;
Qy 580 ACTGACGGTATGAAAGCGGAGCGGTATATTTCTCATCGGAAAGCGCAAGTCACTTC 639
Db 5746 ACTGATTAACCGGGGTAGTGGGTATATTTCTGACGGACACAGGCCCTGGACATTCG 5805
Qy 640 AACGAAATACATGATCGTCAATGTAATCAAAACCTATATGATCGGCGCTCATG 699
Db 5806 AAAAGAAAGTCGTTCTGCAAAACAGTTTCAAAACCGACCTTGAGGCGCAATGCTTG 5865
Qy 700 AGAAGTATTAAGCGCCCGCGCTGATCTGAAAGAGAAATGTTACAGAAAACTGC 759
Db 5866 AAAAGATTAATGCTCCCGCTGCTGACAGCTGAAAGAGAAACAACTCAATCAGTACC 5925
Qy 760 AATTATGCGCTCTGAAAGAAATAGAAAGAGTATCAATCAGAAAGTAAATATGA 819
Db 5926 AGATGATGCCACCGAAGCCAAACAAAGTAGTACAGTCTCGTAAAGTAGAAATCAGA 5985
Qy 820 AAGCAATTACAGCGGAGCACTATTCTGAGATTGGGACATATCTATCATCAGAACTGA 879
Db 5986 AAGCCATTAACCACTGAGCGCACTACTGACAGACTAGCACTGTATTAACCTGCGACAG 6042
Qy 880 ATCCGTGAGTGTTCAGAGTCAATTAATCTGTACCAATCTACTGCAACGTTAATTA 939
Db 6043 ATCAAGCCAGAAATCTATTAAGATCACTATTCGAAACCTATTGTACTCAGTAGGTACCG 6102
Qy 940 ACAAGTTTACATGTGACAGAGTGCCTGTTAAACGTGCAACTTATGTTATTCAGAGAAAT 999
Db 6103 CGAATTAACCTCGATCCACAGTCACTGCTGTAGCTGTCTGTAAACAATATCTGATGAGAACT 6162
Qy 1000 ACCCTACAGTACGCAATTTATTAACAGATTAATCAATGAGTATCTGATGATGTTG 1059
Db 6163 ATCCGACAGTAGATCTTATCAATTAAGTACAGATGATGATGATGATGATGATGATG 6222
Qy 1060 ACGGCGCATGCTGCTGTATGATACAGCCACTTTTGTCCGGCTAAACTGAGAACTACC 1119
Db 6223 ACGGACAGTGCCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 6282
Qy 1120 CAAAGAAAGCATGATTTTGTGACGCGCAAGATTAAGATGAGCTGCTCATACAGA 1179
Db 6283 CGAAGAAAGCATGATTTTGTGACGCGCGCAAGATTAAGATGAGCTGCTCATACAGA 6342
Qy 1180 ATACATTAACAAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1239
Db 6343 ACAAGCTACAAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6402
Qy 1240 GAGATTAACCTGCTTATGATTTGCGCGGCAATTTATGTTGTTTCAAGAAATACGAT 1299
Db 6403 GTGAACTGCGCAACACTGACTGACGCAATTCATATGCAATCTTTCGAAATATGAT 6462
Qy 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
Db 6463 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6522
Qy 1360 TTATGCAATTTGTAAGAAAGCTGAAAGGCGGCAACAGAGATGTTTCCAAATATCTC 1419
Db 6523 TCACCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6582
Qy 1420 ATATATCTAAACCGTTGCGAGAGATGATGATGATGATGATGATGATGATGATGATG 1479
Db 6583 ATATATTTGCTCCATTTGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 6642
Qy 1480 ATGTCAAAAGTTACTCCCGGACGAAACATACAGAGAGCGGCTTAAAGTGCAGATTTC 1539
Db 6643 ACGTGAAGTTTACACGAGGACGAAACACAGAAAGAAAGTGAAGTGAAGTGAAGTGA 6702
Qy 1540 AGGCTGCAATCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1599
Db 6703 AAGCCGAGAAACCTGCGAGCTGCTTACTTATGCGGAGTTCACCGGAAATTAAGTGCCTA 6762

QY 1600 GACTGATATGCGGTCTTTCGCCAAATATCATATCTCTTGACATGTGACGGGAAGTT 1659
Db 6763 GCGTTACCGCGCTTGCTTCCAAACATTCACACCTTTTGACATGTGCGGGAGATT 6822
QY 1660 TTGATGCATTAATTCCTGAACATTTCCACACGGGGACCCAGTATTGGAAAGGACATCG 1719
Db 6823 TTGATGCATTAATTCCTGAACATTTCCACACGGGGACCCAGTATTGGAGAGGATATCG 6882
QY 1720 CGTCTGTTGATTAAGCGAAGACGACGCTATCGCCATTTCCGCGTTGATGATCTTTGAGG 1779
Db 6883 CATCATTTGACAAAAGCCAAAGACGCTATGCGCTTAACCGGTGTGATGATCTTTGAGG 6942
QY 1780 ACTTAGTGTGACCAACCGCTCTTAGATTGATAGAGCGCGCTTGCGCAATATCAAT 1839
Db 6943 ACTTGGGTGTGATCAACCACTACTGACTGATGAGTGCCTTTGGAGAAATATCAT 7002
QY 1840 CTGTGCACCTACTATACAGAAAGAGGTTAAATTTGGTCCATGATGAAATCCGTATGT 1899
Db 7003 CCAACCATCTACATGAGGTACTCGTTTAAATTCGGGGCGATGATGAAATCCGGAATGT 7062
QY 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTGCTGACAGACTACGTTG 1959
Db 7063 TCTTCAACATTTTGTCAACACAGATTGATGTGTTATGCGCAGCAGAGTACTAGAG 7122
QY 1960 AACGGTTAACACGTCAGCGTGCAGCGCTCTATCGCGACGATTAACATAGTGCATGTG 2019
Db 7123 AGCGGCTTAAACGTTCCAGATGTGCAGCGTTCATTGGGAGACAAACATCATATGAG 7182
QY 2020 TCGTCTCCGACACCTTGATGGCGAGAGATGGCCACTTGGCTGAACATGGAATTAAGA 2079
Db 7183 TAGATCTGACAAAGAAATGCTGAGAGGTGCGCACCTGCTCAACATGAGGTTAAGA 7242
QY 2080 TTATGATGACGTTATGTTGATCAAGCACCTACTCTGTTGGGGGATTTATCCTGTTGG 2139
Db 7243 TCATGAGCGCATCTCGGTGAGACACACTTACTCTGCGCGGATTTATCTTGACAG 7302
QY 2140 ACCAGATTAACAGGACAGCTGACAGTGCAGACCTCTTAAGAAAGGCTTTTAAGCTTG 2199
Db 7303 ATTCGTTACTTCCACAGGTGCGCGTGGCGGACCCCTGAAGAGGCTTTAAGTTGG 7362
QY 2200 GAAAACCATTTGCGATCGATGATACCCAAAGACTGCGACCGCGCGGACCTGCATGATG 2259
Db 7363 GTAAACCCCTCCACCCGACGACGAGCAAGAAAGAAAGCGCCCTGCTGATGATG 7422
QY 2260 AAGCATGCGATGGAACAGAAATTTAGGAGAGATTGATGAAAGCCGCGAGAAATCA 2319
Db 7423 AAACAAAGCGGTGTTAGAGTATTAACAGSACTTTAGCAGTGGCGGAGACGCC 7482
QY 2320 GATACGATCATATCTGCGAGGCTGATCATCGTCTCTGTCACGTTAGCGGAAAGCG 2379
Db 7483 GGTATGAGGTAGACAAATTTACACTGTCTACTGSCATTTAGAACTTTTGCCGAGACA 7542
QY 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGCCCAATCACTCTTAGCGCTGACCTTAATAG 2439
Db 7543 AAAGGCAATTCACACCATACAGAGGGGAAATTAACATCTTACCGTGTCTTAATAGT 7602
QY 2440 TGAAGTAGACAGCACTAC-----CCACGGGAGAAATGTTTCCATACC 2486
Db 7603 CAGCATAGTACATTTCAATCTGCTAATACATAACAACAACAACATGATAGAGATTCT 7662
QY 2487 CTCAGCTGAACCTTTCACCAAGTTTACCTTACAAATCCGATGGCTTACCGAGATCCAAAC 2546
Db 7663 TTAAATGCTGTGGCGCGCGCCCTTCCGGCCCCCAGTGCATGTGAGAGCGCGGAGAA 7722
QY 2547 CTCCTAGGCGCGCTGAGAGCGGCTTTGCGCCCCCGCTGCTGCTCAAAATGCAAGATCTTA 2606
Db 7723 GAGGCGAGCGGCGCGCGATGCTGCGCCCAACGGGCTGCTTCAAAATCCAGCACTGA 7782
QY 2607 GGAGGTGATATGATCACTTGAATTTCAAAACAACATCACTTAATCCGCGCCAGTCCA- 2665
Db 7783 CCACAGCGGTACGTCCCTAGTCAATTTGACAGGCAACTAGACCTTAACCCCTACGCTCAC 7842
QY 2666 -----CCGCAAAAGAAAGAGTGTCTCTTAAGCCAAACCTACTAGCTTAAGAAAGA 2720

Db 7843 GCCGCCACCGCGCAGAAAGACAGCGGCCCAAGCAACCAAGCCGAAGAAACCA 7902
QY 2721 AGAAGCAGAACCCAGAGAGACGAAC--GCAGCTTAATCCAGGAAAGACAACTA 2777
Db 7903 AAAGCAGAAAGAAAGAAAGCAACTGCAAAACCCGAAGAAAGACAGCGCA 7962
QY 2778 TGTGATGAAAGTGGAGTCGACAAAGACATTTCCGATCATG---CTGAACGGCCAAAGTGA 2834
Db 7963 TGGCATTAATTTGAGAGCCACAGATGTTTGACGTCAAGAAAGAGAGCGAGTGTCA 8022
QY 2835 ATGATATGCTGCTGCTGTGCGAAGAGCTGATGAACCACTCCACGTTAAGGAAAA 2894
Db 8023 TCGGCGACGCACTGCGCATGGAAGAAAGTAAATGAACCTTGCACTGAAGAAACCA 8082
QY 2895 TTGATTAAGACATTAAGCGCCCTGGAATTTGAAGAGGCTATGACATGATGATGAGT 2954
Db 8083 TCGACCACTGTGTCTATCAAGCTCAAAATTTACCAAGTGTCAAGCATAGACATGAGT 8142
QY 2955 ACGGCGACGTTCCCGAGAACATGAATCAGACAGCTGCAGTACACGACGCAAAACAC 3014
Db 8143 TCGCACAGTTCCAGTCAACATGAAAGTGAAGCATTTACCTTACACAGTGAACCCCG 8202
QY 3015 CGGCTTCTACACTGCGACACCGCGCAGTCCAGTATGAGAAATGGAATTTACCGTAC 3074
Db 8203 AAGATTTCTATATCTGCGACACCGAGCGGTGAGTATGAGGTGATTTACATCC 8262
QY 3075 CGAAGAGTGTGGGGGAAAGCGACACGCGGAACCGATCTTGCAACAACAGGCAAG 3134
Db 8263 CTCGCGAGTGTGAGGCGAGAGAGACAGCGGTCTCTCATATGATGATTAATCTCCGTCGG 8322
QY 3135 TTGTGCTAATTTGTTAGAGATGTCAATGAGGCGAGCGTACCGCGCTTTCAGTGTCA 3194
Db 8323 TTGTGCGATAGTCTCGGTGAGTGAATGAAGAAACGAATGCTCTTTGCTGTCA 8382
QY 3195 CTGGAACAGAAAGGGTGAACATTAGGATTAACCCCGAAGTTCGAAACCTGCT-- 3251
Db 8383 CCGGAATATGAAGGAAAGCAATTAAGACGACCCGGAAGGGAAGAGAGTGTGCG 8442
QY 3252 -----CACTAGTTACAGCGCTATGCGTCTTTGAAATGTCACTTTCCATGCGCAAAAC 3305
Db 8443 CAGACCACTGTGTACCGCAATGTGTTGCTCGGAAATGTAGCTTCCATGCGACCC 8502
QY 3306 CACCGGTGTCTATTCCTGACGCGCAAGAACACTGAGTGTGCAAGAGACGTCG 3365
Db 8503 CGCCCATGTCTATACCGCGAACCTTCCAGAGCTTGCAATCTTTGAAGAAACGTA 8562
QY 3366 ACAATCCAAATTTAGACACAGCTGTGAGAAAGTCTTGAATGTCTATCAC---GCCGC 3422
Db 8563 ACATGAGGCTTACGATATCCCTGCTCAATGCAATATGCGGTGGGATCGTCTGCGAAG 8622
QY 3423 CCAAAAGAGATTACCGATATCTTACATGACCAAGTCTCTTACCTGGGCTTCTGCGCT 3482
Db 8623 GCAAAAGAGGTACGATGACATTTTACCTGACAGCCCTTACTTGGGACATGCTGCT 8682
QY 3483 ATTGACGACCTCAACGCGGTGTTACGCCCAATTAATAATGAGAAAGTGTGGAGCAT 3542
Db 8683 ACTGCACATATGAAACCGTGTTCAGCCTGTTAGATGAGAGAGTCTGGGACGAG 8742
QY 3543 CTGATGATGATGATTAAGATTCAGGCTTTCGACCAATTCGCTTACATCAAGGACCA 3602
Db 8743 CGAGAGATTAACCATAGCATACAGACTTCCGCGCAATTTGATAGACCAAGAGCGAG 8802
QY 3603 CTGCGATGTCACAAAATTCGTTATATGCTTTTGACCAAGCAACATCAAGAG 3662
Db 8803 CAGCAAGGCGCAAAAGATACGCTATAGTGTCTTAAGCAAGATACACCTTTAAAGAG 8862
QY 3663 ACAGTATGAGAAATAGCTATCAGCATGTGAACCTGCGCTGTGCTTGGCCAGAAAG 3722
Db 8863 GCACCATGATGATCAATGAAGATTAGACCTAGAGCGGTGAGAGGCTTAGTACAAAG 8922
QY 3723 GGTACTTCTGTGATCAATGTCTCTCAGGTGACAGTGAACCGTCAATCAAGAGCG 3782

Dh 8923 GATACCTTCTCTCGCAAAATGCCCTCCAGGGGACAGCGTAAACGTTAGCATAGTAGTGA 8982
Qy 3783 GAGCATCTGAGAAATTCATGACCGGTGAGAAAAAGATCAGAGAAAGTTGTCCGTGAG 3842
Db 8983 GCAACTCAGCAACGTCATGTACACTGCGCCGCAAGATTAACCAAAATTCGTGGAGCGG 9042
Qy 3843 AGAGATCTTGTCTCCACCCGTCATGGAAGTGTAAAGTGCACGTTAACATCACT 3902
Db 9043 AAAAATATGATCTACCTCCGTTTACGCTTAAAAAATTCCTTGCACAGTGTACACCGTC 9102
Qy 3903 TGAAGAGAGCTGTCGGGTGATCAATACCATGCAAGCGCACGCGCTATTAAGT 3962
Db 9103 TGAAGAAACAACTGACAGGCTACATCACTATGCAAGCGCGGACCCGACCTTATACAT 9162
Qy 3963 CCTATCTGAGAGAGCGTCAGCGCAAGGTACATTTAAACCACTTGGGCAAGAACGCA 4022
Db 9163 CTATCTGGAAGAAATCATCAAGGAAAGTTTACGAAAGCCGCACTGGGAAAGAACATTA 9222
Qy 4023 CCTACGAATGTAAAGTGTGACACTACAGCAGGTAATGAGACGCGAAAGAAATGA 4082
Db 9223 CGTATGAGTGAAGTGGCGGCACTACAGACCAAGACCGTTTGCACCGGCAACGAAATCA 9282
Qy 4083 ACGGCTGACTAAAGCAAAACAGTGCATTGCTTCAAGAGCGCAACGAAATGGGTCT 4142
Db 9283 CTGGTTGACCGCCATCAAGAGATGCGTCCCTTAAAGAGCGACCAACGAAATGGGTCT 9342
Qy 4143 TCAACTGCGCGGATCTTATTTAGGCACACAGACCTCAGTGCAGAGTTAAATTTGCACATTC 4202
Db 9343 TCAACTCAACCGGACTTGTATGACAGATGACACACACCGCCCAAGGAAATTTGATTTGC 9402
Qy 4203 CATTCGCGTGAACCGACAGTCTGCGGTTCCGTTAGCTACACGCGCTACAGTCAAG 4262
Db 9403 CTTTCAAGTTATCCCGAGTACTGTACATGTCCTGTTGCCAGCGCGGATGTAATAC 9462
Qy 4263 AGTGTCTAAAGGATCAACCTCCACTGACTGCAATGCGAACAACTTGTGCAACGA 4322
Db 9463 ATGCTTTAAACATACATCGCTCCCAATTAGATACGACCACTTGACATTTGTCAACCA 9522
Qy 4323 GAAATTTGGGCTCGAGACACGCAACGCAAGAAATGATTAACAGGCTTACATTCAGGA 4382
Db 9523 GGAACATGAGGGGCAAAACCGGAAACCACTGATGATCGTGGAAAGCGGTCAAGAA 9582
Qy 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTACGTATGGGTAAACATTAACAGTCAAG 4442
Db 9583 ACTTCAACGCTGACCGAGATGGCTTGAAATCATATGGGAAATCATAGCCAGTGAAGG 9642
Qy 4443 TCTGGGCCAGAGTGGGACCAAGCGCAACCAATGATGCGCGATGATCATATCC 4502
Db 9643 TCTATGCCAAGAGTCAACACGAGAGACCTTCAACGATGGCCACAAGAAATAGTACAGC 9702
Qy 4503 ACTATTAATCGGATTCAGTCTACATGTCATTTGTGCTGTGTGTGCTGCTTGTGTA 4562
Db 9703 ATTAATACATCGGCTCTGTGTACACATCTTAAGCGGTGATCAACGTCACGTTGGGGA 9762
Qy 4563 TCTGTGTAGGACATGACATGACAGCTTGACATGCGCAAGAAAGAGATGCTGTA 4622
Db 9763 TGAATATGGGTATACCTTGTGACAGTGTATGTGCTGTAAAGCGCGGTGATGCTGTA 9822
Qy 4623 CGGCATAAGCGCTTGGACCGAACGCAAGGTAACCCACAGATTAAGCGGTTTGTGCTGA 4682
Db 9823 CGGCATTAAGCGCTTGGACCGAACGCGTAATCCCAATTCGCTGGCATCTTTGTGCTGG 9882
Qy 4683 TTGGCCCAACCAAGCTTGAACATTTGAGAAATTTGACATCTGTGTGTTTAAACAC 4742
Db 9883 TTAGGTGCGCAATGCTAAACGTTTACCGAGACATGATTAATTTGTGTGCAACATC 9942
Qy 4743 AACGTTTCTGCGGACAGTTGTGATTCCTCTGAGAGCGCTGTATTTCTGTTCCGCT 4802
Db 9943 AGCGTTTCTTGTGATCAAGTTGTGATACCTTTTGGCGCTTTTATCGTTTCTAATGCCCT 10002
Qy 4803 GCTTTTCAATGCTGATGCTTTTATTTATTTGTTGACAGCGCTGTGCTGGGAAAGGTAGCG 4862
Db 10003 GCTGCTCTGCTGCTGCTTTTATTTATTTGTTGCTGGCGCTTACTGCGAAGGTAGCG 10062

Qy 4863 CTTTCAACATGCGACACTGTGCAAAATGTTCCGGGATCCGCTATTAAGCGGTGCTG 4922
Db 10063 CTTAGAACATGCGACACTGTGCAAAATGTTCCGGGATCCGCTATTAAGCGGTGCTG 10122
Qy 4923 AACGCGAGTTACGGGCACTTAACCTGAGATCAAGCTGTCTGATCGAAATTAAC 4982
Db 10123 AAAGGAGGATGAGCGGCTCAATTTGAGATCACTGTCAATGCTCGGAGGTTTGC 10182
Qy 4983 CTTCACTAACAGAGATGATGATCTGTGCAAAATTTCCACAGTCACTTCTTCCACAG 5042
Db 10183 CTTTCAACCAACAAAGATTAATTAATCTGCAAAATTCACACTGTGTCTTCCCAAAA 10242
Qy 5043 TTAATGCTCGGGTCCCTGAGTGAAGGATCTTCAAGGCGGATTAACATGCGCG 5102
Db 10243 TCAAATGCTGCGCTCTTGTGAATGTACAGCGGCTTCAATGAGATTAATCTGCAAG 10302
Qy 5103 TTTTGGCGGTGTATCCCTTTCATGTGGGAGCGCAATGCTTCTGTACAGTGA 5162
Db 10303 TCTTGGAGGGGTCTACCTCTTATGTGGGAGAGGCAATGTTTGGACAGTGA 10362
Qy 5163 ACACACAATGATGAGGCTTACGTGATGCTCTCAAGCTGCACTATATGATCAAGCAG 5222
Db 10363 ACAGCCAGATGATGAGGCGTACGTGCAACTGTCAAGATTCGCGTCTGACACGCGC 10422
Qy 5223 TCGCACTAAAGTTCAACAGCTGCTGTGAAGTGGGCGCGGTATGATTAAGGCAACA 5282
Db 10423 AGCGATTAAGTGTGACACTGCGCGCATGAAATGAGATCTGCTATGTATAGGAAACA 10482
Qy 5283 CCAACCGGCACTGATACGTTTGTCAATGCGCTCAAGCGAGTTCCTCAAGGACCTGA 5342
Db 10483 CTACAGATTTCTTAAATGTGTACGTGAAGAGTCAACACGAGAACGCTTAAAGACTTGA 10542
Qy 5343 AGTCAATGAGAGGCGCATATCAAGCGCTTTTACCTTTGACATTAAGTGTATCA 5402
Db 10543 AAGTCAATGCTGAGCAATTAAGCATGTTTACGCGCATTCATCAATTAAGTGTATTC 10602
Qy 5403 GAAAGGGCTTGTAAACATACAGACTTCCCTGAGTATGAGGCTATGAACCAAGAGCGT 5462
Db 10603 ATGCGGCTGTGTATACATTAATGACTTCCGGAATATGAGAGATTAACAGAGCGT 10662
Qy 5463 TCGGCAATTAACAGACTCTGCTGATGATCTACAGATAGTACCCGCACTGACATAC 5522
Db 10663 TTGAGACATTTCAAGTACTCTCTTGTACAGAGATCTCATGCGCAGACAGACTTGA 10722
Qy 5523 GGTCTGAAAGCTTCTGTAAAGAAATCAAGTCCCTTCAACCCAGCAGTATCAGGCT 5582
Db 10723 GGTACTCAAGCTTTCGCAAGAAAGTGTATGCTGTCAGCGAGCGCATCAAGAT 10782
Qy 5583 ATGAATGTGAAGAAACATCTCAGAGACGACCCCTGCAAGAAACAGCAATTTGATGTA 5642
Db 10783 TTGAGATGTGAAGAAACATCTCAGAGCGGCACTGAGAGAAACGCACTTTGGGTGTA 10842
Qy 5643 AAATTAAGTGAAGCTCTGCGAGCGTCTAATGTGCTTACGGGCAATCTTATCTGA 5702
Db 10843 AGATTCAGTAATATCCGCTCCGAGCGGTGAGCTGTATACGGGAACATTTCCATTTCTA 10902
Qy 5703 TTGACATCCCTGATGACGCTTTTGTGATCATCAAGATCAACCAAAATTTTAAAGTTA 5762
Db 10903 TTGACATCCCGAAGCGCTCTTATACAGACATCAGATGACCACTGCTCAACATGTA 10962
Qy 5763 GCTGCAAGTGAAGACATGATTAATCTGACAGCTTGTGTGCTTCTAATTAACAT 5822
Db 10963 AATGTAAGTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11022
Qy 5823 ACAAGCTGACAGGAGAGCAATTTGTCAGATTAATCTCCATCCACAGACAGCTGTTTGA 5882
Db 11023 ATGATTCGACCGGAGAGTCAATGCGGTATATTCGATTCGAGAGACAGCAACTTCC 11082
Qy 5883 AGGAGGACACACATGATGATGCTGCTGAGCAGATTAACATTAATTTAGCAGATGTA 5942
Db 11083 AAGATGCAAGTATCATGATCTCTGAGAAAGAGAGCGGTGACAGTATTAATTAAGACCGCGA 11142

QY 5943 GCCCAGCAAAATTTTATAGTTTGGCTATGCGGAGAGTCCACTGCTGATGCTGAT 6002
DB 11143 GTCCACAGCGCACTTTATGATGCTGTGTGGAGAGCAACATGCAATGAGAAAT 11202
QY 6003 GTAAACCAACCGGCGGACCACTAATTGGAGAACCACTAAGTGCACCAAGATTCCAG 6062
DB 11203 GTAAACCAACCGTACCATATGCTGAGACACCCACAAAATATACCAAGATTTCAG 11262
QY 6063 CGGCACTTCCAAAACATCTTGAACCTGGCTGCTGCACTGTTTGGGGAGATCATCC 6122
DB 11263 CGGCATCTCAAAAACATCATGAGTGGCTGTTGCTTTGCGCGCGCTGCTGCG 11322
QY 6123 TCATTGTTAGAGCTTATAGTGTGCTGCTGAGCTTATGCTTAAACACCTAGAT 6182
DB 11323 TATTATTATAGACTTATGATTTTGTGCTGAGCATGATGCTGACTGACAGAAAT 11382
QY 6183 GACTGAGCGGCGGACACTGA-CATAGCGTTAAACTGATGCTTCCGAGAGGCTGCT 6241
DB 11383 GACCGCTACGCGCCCAATGATCGACACGAAACTGATGCTCCGAGAACTGATGT 11442
QY 6242 GCATATGCGCAGCGCGCT 6261
DB 11443 GCATATGCTATGAGCTGCT 11462

RESULT 7
AAC64507
ID AAC64507 standard; DNA; 11703 BP.
XX AAC64507;
AC
XX
DT 14-FEB-2001 (first entry)
XX
DE Alphavirus SinchironLP virus nucleotide sequence.
XX
KW Alphavirus; Sindchiron virus; SinchironLP virus; immune response;
KW infection; human dendritic cell; immunostimulatory; cytostatic;
KW virulence; fungicide; antibacterial; antiparasitic; vaccine; cancer;
KW pathogen; antigen presenting cell; ds.
XX
OS Alphavirus.
XX
PN WO20061772-A2.
XX
PD 19-OCT-2000.
XX
PF 14-APR-2000; 2000WO-US10722.
XX
PR 14-APR-1999; 99US-0129498.
PR 09-AUG-1999; 99US-0148086.
PR 22-MAR-2000; 2000US-0191363.
XX
PA (CHIR) CHIRON CORP.
PI Polo JM, Dubensky TW, Frolov I, Gardner JP, Otten G, Barnett S;
PI Driver DA;
XX
XX WPI; 2000-619231/59.
DR
XX
PT New alphavirus that infects human dendritic cells for use in generating
PT an immune response to pathogenic agents such as bacteria, viruses,
PT fungi, parasites and cancer and for biological assays -
XX
PS Claim 10; Fig 2C; 83pp; English.
XX
XX The present invention describes an isolated alphavirus (AV) which
CC infects human dendritic cells and is not of American Type Culture
CC Collection (ATCC) number VR-2526. AAC64506 and AAC64507 represent the
CC nucleotide sequence of the specifically claimed Sindchiron virus and
CC SinchironLP virus. The new AVs have immunostimulatory, cytostatic,
CC virulence, fungicide, antibacterial and antiparasitic activities and
CC can be used in vaccines. The AVs are used to infect dendritic cells,
CC preferably human cells. A heterologous sequence can be introduced and

CC expressed in human macrophages or antigen presenting cells in vivo and
CC in vitro, for use in biological assays. The AV-based vector systems
CC are used to generate an immune response to cancer or a pathogenic
CC agent, such as, bacteria, fungi, parasites or viruses. The AV can be
CC used to infect human dendritic cells; macrophages or antigen
CC presenting cells that previously could not be infected using an AV or
CC AV variant. The AV vectors are targeted directly to antigen presenting
CC cells.
SQ Sequence 11703 BP; 3308 A; 3050 C; 2911 G; 2434 T; 0 other;
Query Match 36.7%; Score 2382; DB 21; Length 11703;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3694; Conservative 0; Mismatches 1985; Indels 41; Gaps 8;
QY 580 ACTGACGATGAAAGCGGAGACGATATATTTCTCATCGGAAACAGGCCAGTCACTTC 639
DB 5746 ACTGACTAACCGGGGTAGTGGTACATATTTTGGACGACACAGCCCTGGCACTTGC 5805
QY 640 AACGAAATCAGTACGTCATGTAACTACAAAGAACCTATATTGGATGGGCCCTGCATG 699
DB 5806 AAAAGAGTCCGTTCTGAGAACCACTTACAGAACGACCTTGGAGCTGATGCTCGG 5865
QY 700 AGAAGTATTAAGCCCGCGCTGATCTGAAAGAGAAATGTTACAGAAAGAACTGC 759
DB 5866 AAAAGATTCATGCCCGGTGCTGACAGCTGAAAGAGAACCACTAAACTAGTACC 5925
QY 760 AATTATGCGCTCTGAAAGAAATAGAACAGATCAATCAACGAAAGTAAATATGA 819
DB 5926 AGATGATGCCCAACGAGCCAAACAAAGTATGATCACTCTGTAATAGAAATATCGA 5985
QY 820 AAGCAATTACAGCGGAGGAGCTCATTTCTGATTGGGACATATCTATCATCAGAGCA 879
DB 5986 AAGCATACCACTGAGGACTTACTGTAGACTACGACTATATCTTCTGACAG--- 6042
QY 880 ATCTGTGAGTGTATACAGATCAATTATCTGTACCAATCTACTCTGCAACGATATTA 939
DB 6043 ATCAGCCAGATGCTATATGATATCACTATCGAAACATGTTACTCCAGTACGATCCG 6102
QY 940 ACAGTTTACATCTGCAGAGGTCCGGTTAAAGCTGCACTTACTTATCCAGAGATT 999
DB 6103 CGAACTACTCCGATCCACAGTTCCTGTGCTGTATCAACTATCTGATGAGAACT 6162
QY 1000 ACCCTACAGTACGAGTATTTGTAACAGATGATAGATGCGATCTTACAGATGGG 1059
DB 6163 ATCCGACATGACATCTTATCAGATTTACTGACGATGATGCTTACTGATATGGTAG 6222
QY 1060 ACGCGCATCGTGTCTGTATACAGCCACTTTTGTCCGGCTAAACCTGAGAACTACC 1119
DB 6223 ACGGACAGTGCCTGCTGATGATCTGACACTTCTGCCCCGCTAAGCTTGAAGTTACC 6282
QY 1120 CAAGAGCATAGCTATTTTGCAGCGAGATAGATACGCCGCTCCATGCTTATACGA 1179
DB 6283 CGAAAAAATAGATATAGAGCCCGCAATATCCGACAGTCCGTTCAATCAGATGAGA 6342
QY 1180 ATACATTACAAATGATATGCTGCTGACGCTATTAAGAAATTTGAACCTTACCAATGC 1239
DB 6343 ACAGCTACCAAAATGTCTCATTTCCCGCACTAAAGAAATTTGCAACCTACGAGATGC 6402
QY 1240 GAGAAATACCTGTCTTGAATCGCGGCACTTAAATGTTGTTTCAAGAAATACGAT 1299
DB 6403 GTGAACCTCCCAACTGACACTCAGCAGACATTCATATGCTGATGCTTGAAGAAATGAT 6462
QY 1300 GCAATGATAGTACTGGGATACCTTTGCGATTAACCTATTGCGCTAATCAAGAAAG 1359
DB 6463 GTAAATGAGATATTTGGAGAGATTCGCTCGAAGCCATATAGATTAACCTAGATTGG 6522
QY 1360 TTAGCAATATGATCAAGAGTGAAGGCGGAAAGCAAGCACTGTTTGGAAATCTC 1419
DB 6523 TCACCGCATATGATGCTGATGATGAAGGCTTAAAGCCGCGCACTATTTCGAAGAGCT 6582
QY 1420 ATATCTTAAACCGTTGAGAGATACCAATGATCAATTCGATGATTAAGAGAG 1479

Dh 6583 ATAATTGTCCTCATGCAAGAGTGCCTATGATGATTCGTGATGACATGAAGAAAGG 6642
Qy 1480 ATGTCAAGTTACTCCGGGACGAAACCTTAGAGAGCGGGCTTAAGTGAAGTTATTC 1539
Db 6643 ACGGAAAGTTACACGAGGACGAAACACAGAAAGAAAGACCAAAAGTAAAGTGTATAC 6702
Qy 1540 AGGCTGAGATCCCTCTGCTACCGCTTACCTTTGCGGGATCCATCGGGAAATTAGTCGTA 1599
Db 6703 AAGCGGAGAACCCCTGGCGACTGCTTACTTATGCGGATTCACCGGAAATTAAGTCGTA 6762
Qy 1600 GACTGAATGCGGTCTTTCGCAAAATATCATCTCTTTCGACATGTCAAGCGGAAGATT 1659
Db 6763 GCGTTAAGCGCGCTTGTCTTCAAAATCATACGCTTTTGAATGTGCGCGGAGATT 6822
Qy 1660 TTGATGCGATTATGCTGAACATTTCCACGCGGACCGCATTTGGAAACGACATCG 1719
Db 6823 TTGATGCAATCATGACGAAACACTTCAAGCAAGGCGACCCGGTACTGGAGCGATATCG 6882
Qy 1720 CGTGTGTTGATTAAGGCAAGACGACGCTATCGCATTTGCGCGTTGATTCCTTGAG 1779
Db 6883 CATCATTCGACAAAAGCCAGAACGACCTATGCGGTTAACCGGTCTGATGATCTTGGAG 6942
Qy 1780 ACTTAGGTGACCAACCGCTCTTAATTTGATAGAGCGCGCTTCGCAATATCAT 1839
Db 6943 ACTGGGTGTGGATCAACCACTACTCGACTTGATCGAGTGGCGCTTGGAAATATCAT 7002
Qy 1840 CTGTCGACCTTACCAAGGAAACGAGTTAAATTTGGTCCATGATGAAATCCGGTATGT 1899
Db 7003 CCACCCATCTACCAAGGGTACTCGTTTAAATTCGCGGCGATATGAATCCGAAATGT 7062
Qy 1900 TCTTAAAGCTGTTGTCAACACACTAGTCATATCATGATTTGCTAGCAGATACATCGTG 1959
Db 7063 TCTTCAACATTTTGTCAACACATTTGAAATGTGTTATGCTCCAGCAGAGTACTAGAG 7122
Qy 1960 AACGCTTAAACGCTGACGCGTGCAGCGCTCTATCGGAGCATPAACATAGTGTG 2019
Db 7123 AGCGGCTTAAACGCTGAGATGTGACGCTTCAATTTGGGAGCAACATCATATGAG 7182
Qy 2020 TCGTCTCCGACACCTTGATGCGGAGAGATGCGGCACTTGGCTGAACATGGAATMAAA 2079
Db 7183 TAGTATCTGACAAAGAAATGCTGAGAGTGGCGCACCTGGCTCAACATGGAAGTTAGA 7242
Qy 2080 TTATTGATGAGTATTGATTAAGGACACCTTCTGTGGGGATTTATCTGATG 2139
Db 7243 TCAATGAGCGATCATCGGTAGAGACCACTTACTTCTGCGCGGATTTATCTTGCAG 7302
Qy 2140 ACCAGATPAACAGGACAGCCTGACAGTGCAGACCCCTCTAAAAAGGCTTTTAAAGCTG 2199
Db 7303 ATTCGGTTACTTCCACAGCGTGCAGCGCTGGCGAACCCCTGAAAAAGGCTGTTAAAGTTGG 7362
Qy 2200 GAAAAACATTGCGATGATATACCAAGACTGAGACCGCGCGGCACTGCATGATG 2259
Db 7363 GTAAACCGCTCCAGCGGACAGAGCAAGAAAGAAAGCGCGCTTGTGTAGATG 7422
Qy 2260 AAGCATGCGATGAAACGAATTTGAATTACGAGAGTATGTAAGGCGCTTGAATCCA 2319
Db 7423 AAACAAAGGCGTGGTTTAAAGTAAAGTAAACAGGCACTTTAGCAGTGGCGGTGACAGCC 7482
Qy 2320 GATACGAGATCATCTGGCAGCGCTGATCATCACGCTCTGTGTCAAGCTTACCGGAAAGCG 2379
Db 7483 GGTATGAGTGAACAATTTACACTGTGCTACTGAGAACTTTTCCAGAGCA 7542
Qy 2380 TTAAAGACTTCAAGAGCTTAGAGAGGACCAATCACTCTTACGGCTGACCTTAAATAG 2439
Db 7543 AAAGACATTCAGGCACTGAGAGGGAATTAAGCATCTCTACGGTGTCTTAAATAGT 7602
Qy 2440 TGAGTATGACAGCGCACTAC-----CCACGGGAGAAATGTTTCATACC 2486
Db 7603 CAGCATTAATCTTCACTCTGACTAATTAATAACACACCAACCAATGAATAGAGATTCT 7662
Qy 2487 CTCAGCTGAACCTTTCACAGCTTAACTTACCAATTCAGATGCTTACCGAGATCCAAAC 2546
Db 7663 TTATCATGCTGCGCGCGCGCTTCCGCGCCCACTGCAATGTGAGGCGCGGAGAA 7722

Qy 2547 CTCTAGAGCGCGCTGAGGCGGTTTCGCGCCCGCTGCTGCTCAATGAGATCTTA 2606
Db 7723 GAGAGAGCGCGCCGATGCTGCTCCCGACAGCGGCTGCTTCTCAATCTCAAGACTGA 7782
Qy 2607 GAGGTGATATGATCACTTGAATTTCAACAGATCACTTAATCCGCGCGGCTGCA- 2665
Db 7783 CCAAGCGGTGATGCTGCTGATGAGCAGGCACTAAGCTCAACCCCGACGTCTAC 7842
Qy 2666 -----CCGCAAGAAAGAGAGAGTGTCTTAAAGCCAAACCTTCTGACCTTAAAGA 2720
Db 7843 GCGCGCACCGCGCGAGAGAGCAGCGCGCCAGAGCAACCAAGGAGCCAGAAACCA 7902
Qy 2721 AGAAGCAGAGCAAGCAAGAGAGCAAAAC--GCAAGCTTAAACAGGAGAAAGCAACGTA 2777
Db 7903 AAACGAGAGAAAGAAAGAAAGCAACCTGCAAAATCCAAACCCGAAAGACAGCGCA 7962
Qy 2778 TGTGTATGAAGTTGAGTCCGACAGACATTTCCGATCATG---CTGACGCGCAAGTGA 2834
Db 7963 TGGCATTTAGTTGAGGCGGACAGATCGTTGACGTCAAGAAAGAGAGATGTCA 8022
Qy 2835 ATGATATGCTGCTGCTGTTGTGAGAGAGCTGATGAACCACTCCACGTTGAAGAAAA 2894
Db 8023 TCGGCGACGCACTGGCCATGAGAAAGGTAATGAACCTTGCACGTGAAAGAAACCA 8082
Qy 2895 TTGATATGACCAATTAGCGCGCGTGAATTTGAAGAGGCTAGCATGTAGACTTGGAGT 2954
Db 8083 TCGACACCTGTGCTTCAAAAGCTCAAAATTTCAAGTGTGACATGACATGAGAT 8142
Qy 2955 ACGGCGAGCTTCCCGAAGCATGAAATCAGACAGCTGCAGTACACGAGGACAAACAC 3014
Db 8143 TCGACAGTGTCCAGTCAATGAGATGAGGACATTCACCTTACCAAGTGAACACCCG 8202
Qy 3015 CGGCTTCTTCAACTGGCACACGCGGCACTCAATGAGAAATGAGAGTTTACCTGAC 3074
Db 8203 AAGGATCTTAACTGGCACACGAGCGGTGAGTATGAGAGTGAATTTTAAACATCC 8262
Qy 3075 CGAGAGAGTGGCGGGAAAGGCGACAGGAGAAACCGATCTGGAACACAGAGGCGAG 3134
Db 8263 CTGCGAGTGAAGAGGAGAGAGAGACAGCGTGTCTCGATCATGATATCTCCGCTGG 8322
Qy 3135 TTGTGCTATTTGTTTGAAGGTGCAATGAGGAGCGCGTACGCGCTTTCAGTGTCA 3194
Db 8323 TTGTGCGATAGTCTCGTGAAGCTGATGAAGAACAGAACCTGCTTCCGTGTCA 8382
Qy 3195 CTGGAACCAAGAAAGGCTGACATTTAGGATATCCCCGAAGTTTGAACCTGTG--- 3251
Db 8383 CTTGGAATTAAGGAGAGACATTAAGACGACCCCGAAGGAGACAGAAAGTGTGCG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGCGTCTTGAATGACGTTCCATGCGACAAAC 3305
Db 8443 CAGACCTACTGTGACGCGCAATGTGTTGCTGGAATGTAGCTTCCATGCGACCGCC 8502
Qy 3306 CACCGGTGCTATTCTACGACGCGAAGCAACCTGACGCTGCTGGAAGAGAGCTGCG 3365
Db 8503 CGCCACATCTATACCCCGGAACCTTCCAGAGCGCTCGACATCTTGAAGAGAAAGTGA 8562
Qy 3366 ACAATCCAATTTACGACGCTGCTGAGAAAGCTTTGAAATGTCAATC---GCCGCG 3422
Db 8563 ACCATGAGGCTTGAATCACTCTGCAATGCCATTTGCGGTGCGGATCTGTGCGAGAA 8622
Qy 3423 CCAAGCAGATTTACGATGATCTTCACTGACAGCTGACAGCTGCTGAGGTTTGCCTG 3482
Db 8623 GCABAAAGCGTCACTGAGACCTTTTACCTTGAACAGCCCTTACTTGGGACATGCTGCT 8682
Qy 3483 ATTGCAACACTCAACGCGCTGTTTCAAGCCCAATAAATTTGAAGAGCTGTGGAGCGAAT 3542
Db 8683 ACTGCCACATATCGAACGCTGTGAGCGCTGTAAAGATCGAGAGGTCTGGAGACGAAG 8742
Qy 3543 CTGATGATGATGATTAAGATCAAGGTCTGCGCAATTTGCGCTCAATCAGGACGCA 3602
Db 8743 CGAGCATTAACACATATGCAATACAGACTTCCCGCAGTTTGTGATGACAAAGCGGAG 8802

QY 3603 CTGCGGATGTCACCAAAATCCGTATCATGTCTTTCGACCAAGCATGACATCAAGAG 3662
Db 8803 CAGCAAGCGCAAAAGTAACCCGTACATGTCCCTTAAGCAGATGATCAACCGTTAAAGAA 8862
QY 3663 ACAGATAGGAAATAGTATCAGACATCTGAGACCTGCGGTGTCTTGGCCCAAG 3722
Db 8863 GCACCATGATGATCAAGATTTACACCTCAGGACCGGTGTGAAGGCTTACCTCAAG 8922
QY 3723 GGTACTTCCTGTACTCAATGTCCCGAGGTGACAGGTATACCGTACATCAAGAGC 3782
Db 8923 GATCTTCTCTCCGCAAAATGCTCCCTCAAGGGAACGCTAACGGTTAGCATATGAGTA 8982
QY 3783 GAGCATCTGAGAAATTCATGACCCGTGAGAAAAGATCAGAGAAATTTGTCCGTAG 3842
Db 8983 GCAACTGACCAACGTATGTAACCTGGCCGCAATATAACAAATTCGTGGAGCGG 9042
QY 3843 AGGACTCTTGTCCACCCGTCATGGAAGCTGTAAAGTCCACGTTACGATCACT 3902
Db 9043 AAAATATGATCTACCTCCCGTTCAAGTAAAAAATTCCTTGACAGGTGACCGCTC 9102
QY 3903 TGAAGGAGACGCTGCGGGTACATTAACATGACAGGCGACGCGCATTAAGT 3962
Db 9103 TGAAGGAAACAATGACAGGTACATCATATGACAGGCGGAGACCGCATCTTAACAT 9162
QY 3963 CCTATCTGAGGAAAGCTCAGGCGAAGTATCAATTAACCACTTCTGACAGAACTCA 4022
Db 9163 CCTACTGGAAGAAATCATAGGGAAGTTTACGCAAAAGCCGCACTTGGGAAGAACTTA 9222
QY 4023 CCTACGAATGTAAGTGTGCGCATACAGACAGGTAATGTTGACACGCGAACGAAGATGA 4082
Db 9223 CGTATGAGTGCAGAGCGGCGCATCAAGACCAAGACGTTTGCACCGGCAACGAATCA 9282
QY 4083 ACGGTGACATAAGCAAAACGTGATCTTCCCTAAGAGCGAACCAAGAAATGGGTCT 4142
Db 9283 CTGGTGAACCGCCATCAAGCAGTGTGCTTAAAGCGAACCAAGAAATGGGTCT 9342
QY 4143 TCAACTGCGCGGATCTTATAGGCAACAGACACTCAGTGCAGATTAATTCACATTC 4202
Db 9343 TCAACTGACCGGACTTGTATGACATGACGACCAACGCGCCCAAGGAAATTCATTCG 9402
QY 4203 CATTCGCTTGACACGACAGTGTGCGGTTCCTTACGCTCAACGCTTACATGACGA 4262
Db 9403 CTTCACATTTATCCGAGTACTGATGTCTCTGTGCGCACGCGCAATGTAATAC 9462
QY 4263 AGTGTTCAGAGCATCAACCTCCACCTGACCTGACATGACCAACATTTGCTGACAAAGA 4322
Db 9463 ATGGCTTTAAACAATCAGCTCCCAATTAGATAGAACCACTTGACATTTGCTCACACCA 9522
QY 4323 GAAATTTGGGCTGCGAGCAGACGCAACGACAGATGATTAAGGCTCATCATCCAGGA 4382
Db 9523 GGAATCTAGGGGCAAAACCGGAAACCACTGAATGATGCTGGAAGAGCGGTACAGA 9582
QY 4383 ATTTTCTGTGGGGGAGAAAGGCTGAGATGATGAGGTAAACCATGAACCATGACAG 4442
Db 9583 ACTTCAACGTCGACCGAGATGGCTGGAATACATATGAGGAAATCATGAGCCAGTGA 9642
QY 4443 TCTGGGCGCAGAGTCTGCGACCAAGCGACCACTGATGATGCGCATGATCATATCC 4502
Db 9643 TCTATGCGCAAGTACACCAAGGAGACCTTCAAGATGCGCACACCAAAATGATACAGC 9702
QY 4503 ACTATATCATGCGATCAGTCTACATCTGATGATGATGATGATGATGATGATGATGAT 4562
Db 9703 ATTACTACATCGCATCTCTGTATACCATCTTATGCGGTGACATCAGTACCTGAGGA 9762
QY 4563 TCTGTGAGGACATGACATCAAGCAGCTTGACGCAAGCAAGAAAGAGACTGCTGA 4622
Db 9763 TGATGATGCGTAACTGTGACAGTATATGATGCTGTAAGCGGCGGTGAGTCTGA 9822
QY 4623 CGCCATAGCGCTTGACCGAAACGAGGATACCAAGCATTAAGCGGTTTGTGCTGCA 4682
Db 9823 CGCCATAGCGCTTGCGCCCAACGCGATATCCCAATTCGCTGACATCTTGTGCTCG 9882
QY 4683 TTGCGCAACCAACGCTGAACAATTGGAGAAATTTGAACCATCTGTGTTTAAACAC 4742

Db 9883 TTAGTGCAGCAATCTGAACGTTACACGAGAACCATGATGATCTGTGTGAGAACATGTC 9942
QY 4743 AACGTTTCTGTGGGCAACATGTGATCTCTGTGACGCGGTTTATCTGTCCGT 4802
Db 9943 AGCCGTTTCTGTGGTCAAGTGTGATACCTTTGGCGCTTCACTGTCTTAAGCGCT 10002
QY 4803 GCTTTCAAGTGCATGCTTTTATATGTTGATGAGGCGTCTGCGGGAAGTGAAGC 4862
Db 10003 GCTGCTCTGCTGCTGCTTTTATGTTGATGAGGCGCTTACCTGGGGAAGTGAAGC 10062
QY 4863 CTTTGAACATGCGACCACTGTGCCAAATGTTCCGAGGATCCGTTAAGCGTGTG 4922
Db 10063 CCTAGAACATGCGACCACTGTGCCAAATGTTCCAGACATACCTGTTAAGCACTGTG 10122
QY 4923 AACGCGAGTTTACGCGCACTTAACTGAGATCAAGCGTGTCTCATCGAATTAAC 4982
Db 10123 AAAGGCGAGGATATCCCGCTCAATTTGAGATCATGTATGCTCGAGGTTTGC 10182
QY 4983 CTTCACATTAAGAGATGACGTGACCTGCAATTCACACAGTCAATTCCTTACACAG 5042
Db 10183 CTTCACACCAACAGATGATATACCTGCAATTAACAATGTTGCTCCCTCCCAAAA 10242
QY 5043 TTAATGCTGCGGCTCCCTGAGTGCAGAGCATCTCAAGCGGATTAACATGCTCCG 5102
Db 10243 TCAATGCTGCGGCTCTTGAATGTGACCGCGCTTATGACATATACCTGCAAG 10302
QY 5103 TTTTGGCGGTGTACCTTTATGTTGAGGAGCGCAATGCTTCTGTGACATGAGA 5162
Db 10303 TCTTGGAGGGGTCTAACCTTTATGTTGAGGAGCGCAATGTTTGGCAGATGAGA 10362
QY 5163 ACACCAATGATGAGGCGATACGTGAGTTCGTCCAGACTGCACTAATATACAGAG 5222
Db 10363 ACACCAATGATGAGGCGATACGTGAGTTCGTCCAGACTGCACTAATATACAGAG 10422
QY 5223 TCGCATTAAGATTCACACAGTGTCTGAAAGTGGCGTGTATATATACGCAACA 5282
Db 10423 AGCGATTAAGTGTGACACTGCGGAGTGAAGATGACGTGTATGATGATGAGGAACA 10482
QY 5283 CCACCGCCACCTGATACGTTTGTCAATGCGGTCAACCCAGTTCCTACGGAACCTGA 5342
Db 10483 CTACCAATTTCTGTATGTATGATGATGATGATGATGATGATGATGATGATGATGAT 10442
QY 5343 AGTCATGACAGGCGGATATCAAGCGGCTTTTCACTTGAACATTAAGTGTGATCA 5402
Db 10543 AAGTCATGAGGCAATTTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 10602
QY 5403 GAAAGGCGCTGTTTACAACTACGACTTCCGAGATGAGAGCTATGAACCAAGAGCGT 5462
Db 10603 ATCGGCGCTGTGTATCAACTATGATCTTCCGAGATGAGAGCTATGAACCAAGAGCGT 10662
QY 5463 TCGGAGATTTCAAGCATCTCGCTGTATGCTACAGACATATGATGATGATGATGATGAT 5522
Db 10663 TTGAGATCATTAAGTATCTCTTGAATGACAGAGATCTATGCTGACAGACAGCATTTA 10722
QY 5523 GGTCTGTAAGCTTCTGTCAAGACATCAAGTCTCCCTTACACCAAGCATATAGGT 5582
Db 10723 GGTCTGTAAGCTTCTGTCAAGACATCAAGTCTCCCTTACACCAAGCATATAGGT 10782
QY 5583 ATGAATGTAAGAAACAATCAGAGACAGCCCTGCAAGAAACAGACATTTGATGATGA 5642
Db 10783 TTGAGATGTAAGAAACAATCAGAGACAGCCCTGCAAGAAACAGACATTTGATGATGA 10842
QY 5643 AAATTAAGTGAAGCTTGTGAGAGCTTAACTGTGCTTACGAGGACATCTCTATCTGA 5702
Db 10843 AGATGAGTGAATATCCGCTCGAGAGGAGCTGTGCTTACGAGGAAATTTCCATTTCTA 10902
QY 5703 TTGATCTCTGATGAGCTTTTGTGAGATATCAAGATCAACCAATTTTGAAGTTA 5762
Db 10903 TTGATCTCTGATGAGCTTTTGTGAGATATCAAGATCAACCAATTTTGAAGTTA 10962
QY 5763 GCTGACAGTGAAGATGATTTATTTGAGAGCTTTGAGGATTTCTTAACATTAACATG 5822

Db 10963 AATGTGAAGTCAGTAGTGACCTTATTACAGACAGATTGCGGGAGTGAGCCACCTGCAGT 11022
Qy 5823 ACAAGGTGACAGGAGGAGACATTTGTCAGTTCTCCCACTCCAGACAGCTTTTGA 5882
Db 11023 ATGTATCCGACCGGAGGTGAATGCCCCGTACATTGTCATTCAGACAGCAACTCTCC 11082
Qy 5883 AGAAGGACACACATGTCAGTCCGTAGGACATTAACCTACATTTTAGACATCGA 5942
Db 11083 AAGAGTGCAGTACATCTCTGAGAAAGAGGGGTGACAGTACATTTAGCACCCGGA 11142
Qy 5943 GCCCACAAGCAATTTTATGTTTCCGTATGCGCAAGAAATGCACTGCAATGCTGAT 6002
Db 11143 GTCCACAGGCGCACTTATGCTATGCTGTGTGGAAAGACAACATGCAATGCAAGAT 11202
Qy 6003 GTAAACACCGGCGGACCACTTAATTTGAGAACACATTAAGTGCACCAAGATTTCCAG 6062
Db 11203 GTAAACACACAGCTGACCATATTCGTAGACACCCGCAACAAATGACCAAGATTTCCAG 11262
Qy 6063 CGGCAATTTCCAAAACATCTTGAAGTGGCTGCTTGGCACTGTTGGGGAGCATATCCC 6122
Db 11263 CCGGCATCTCAAAAACATCATGAGTGGCTGCTTGGCCCTTTGGGGCGCCCTCTGTCG 11322
Qy 6123 TCATTTGTTAGAGCTTATAGTGTGCTGTCAGCTCTATGCTTATAAACACAGCTAGAT 6182
Db 11323 TATTAATTATGAGCTTATGATTTTGTCTTCACATGATGCTGACTAGACAGAGAT 11382
Qy 6183 GACTGACGCGGAGACATGCA-CATAGCGGTAAACTGATGCTTCCGAGAGAACGTGCT 6241
Db 11383 GACCGCTACGCCCCCAATGATCGACAGCAAACTCGATGATCTTCGAGGAACTGATGT 11442
Qy 6242 GCATTAATGCCACGCGCGCT 6261
Db 11443 GCATTAATGCATCAGGCTGCT 11462

RESULT 8
AAT47668
ID AAT47668 standard; cDNA; 11663 BP.
XX
AC AAT47668;
XX
DT 19-MAY-1997 (first entry)
XX
DE South African Arbovirus 86 RNA transcript cDNA.
XX
KM S.A.A86; Sindbis virus; attenuation; vaccine; diagnosis; antibody;
KM vector; Ockelbo; polyarthritis; ss.
XX
OS South African Arbovirus No. 86.
XX
FH Key location/Qualifiers
FT mutation 215
FT /*tag= a
FT /note= "a guanine for adenine substn. at nucleotide
FT 215 is a preferred silent mutation used in
FT recombinant DNA constructs of the
FT invention"
FT
FT mutation 3863
FT /*tag= b
FT /note= "a guanine for cytosine substn. at
FT nucleotide 3863 is a preferred silent
FT mutation used in recombinant DNA constructs
FT of the invention"
FT
FT mutation 5984
FT /*tag= c
FT /note= "a guanine for adenine substn. at nucleotide
FT 5984 is a preferred silent mutation used in
FT recombinant DNA constructs of the
FT invention"
FT
FT mutation 9113
FT /*tag= d
FT /note= "a cytosine for thymine substn. at
FT nucleotide 9113 is a preferred silent

FT mutation used in recombinant DNA constructs
FT of the invention"
XX
XX MO637220-A1.
XX
XX 28-NOV-1996.
XX
XX 22-MAY-1996; 96MO-US07457.
XX
XX 23-MAY-1995; 95US-0446932.
XX
XX (UNIC-) UNIV NORTH CAROLINA.
PI Davis NL, Johnston RE, Simpson DA;
XX WPI: 1997-020937/02.
XX
XX New recombinant DNA encoding infectious South African Arbovirus 86
PT RNA transcript and related mutants - under control of heterologous
PT promoter; useful in prodn. of attenuated vaccines and to generate
PT diagnostic antibodies
XX
XX Claim 2; Page 13-19; 27pp; English.
PS
XX A cDNA clone (AAT47668) codes for an infectious South African
XX Arbovirus No. 86 (S.A.A86) virus RNA transcript. Novel
CC recombinant DNA constructs comprise the cDNA, pref. mutated to
CC include attenuating and silent mutations, and an upstream
CC heterologous promoter (pref. T3, T7 or Sp6). Infectious
CC attenuated viral particles produced from cells transfected with the
CC RNA transcripts encoded by such cDNA clones are useful in live
CC attenuated vaccines, e.g. against diseases caused by related
CC viruses such as Ockelbo, and can also be used as vectors to express
CC antigens of other viruses. The cDNA clone can be administered to
CC animals to raise antibodies for diagnosis of S.A.A86 infection.
XX
XX Sequence 11663 BP; 3297 A; 3104 C; 2903 G; 2359 T; 0 other;
SQ

Query Match 36.7%; Score 2381.6; DB 18; Length 11663;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 3742; Conservative 0; Mismatches 1994; Indels 74; Gaps 9;
Qy 580 ACTGACCGTATGAAAGCGGAGCGGTATTTTCTCATGGAAGGCGCAAGTCACTTC 639
Db 5707 ACTGCTAACCGGGGTAGTGGGTACATATTTGACGAGACAGCGCCGTGGCATTTGC 5766
Qy 640 AACGAAATACAGTACGTCAATGTAACTACAAAGAACTATTTGATCGGCGCTCATG 699
Db 5767 AAAAGAAAGTCCGTCTTCAGAACCACTTACAGAACCGACCTTGGACCGCAATGTTCTGG 5826
Qy 700 AGAAGTATTTACGCGCGCGCGCTGATCTCGAAAGAGAAATGTTACAGAAAGTGC 759
Db 5827 AAAGAAATCTACGCCCCGCTGCTGACACGTGGAAGGAAACACTAACTCAGGTACC 5886
Qy 760 AATTATCGCCTCTGAAAGAAATAGAAAGCATGATCATACGAAAGTAAATATGA 819
Db 5887 AGATGATGCCACCGGAGAACCAAGAAAGCATGATCATGCGAAAGTAAAGAACAGA 5946
Qy 820 AAGCAATTAACGCGGAGCGACTATTTCTGATTTGGGACATATCTATCATCAGAGTGA 879
Db 5947 AAGCAATTAACGCGGAGCGACTATTTCTGATTTGGGACATATCTATCATCAGAGTGA 879
Qy 880 ATCTGTGAGTGTATTAAGAGTCAATATCTGTAACCAATCTATCTGCAACGTAATTA 939
Db 6004 ATTCAGCAAGATGTATTAAGATCACTTACCGAAACATGTATTCAGAGTATACAG 6063
Qy 940 ACAGTTTACATCTGACAGAGTCCGCTTAAACGTCAACTTATGTTATCCAGAGATT 999
Db 6064 CGAATCTACTGACCCAAAGTTTGTCTAGCTGTTTGAACATATCTGATGAGATT 6123
Qy 1000 ACCCTACAGTACCGATTTATTTATTAACAGATGAATCATGCTATCTTGCATGTGG 1059
Db 6124 ACCGAGGTAGCATTTATTCAGATCACGACGAGTACGATGCTTACTTGATATGTGAG 6183

OY 1060 ACGGCGATCGTCTGTAGATAGAGCCATTTTGTCCGGCTAAACTGAGAGCTACC 1119
DB 6184 ACGGACACGTGCTGCTTGCCTAAGTACTGCAACTTTTGGCCCGCCAAAGCTTAAGATTACC 6243
OY 1120 CAAGAAGCATATGATTTTGGACCGAGATAGATCAGCCGCTCCATCGCTTATACAGA 1179
DB 6244 CGAAAAGACACGAGTATAGAGCCCAACATCCGCACTGCGGTTCATCAGAGATGACAGA 6303
OY 1180 ATACATTCAAAATGTAATTGCTGACGCTAATAAAGATTTGCAACCTTACCAATTC 1239
DB 6304 ACACCTTGAAAACGCTGCTCATTTGCGGCACTAATAAAGAACTGCAACGTCACAAATTC 6363
OY 1240 GAGAAATTCCTGTCTAGATTGGCGGCACTTTAATGTTGATTTGTTCAAGAAATACGAT 1299
DB 6364 GTGAATGCAACACTGAGCTCAGGACACTTCAAGCTTGAATGCTTTGCAAAATTTGCT 6423
OY 1300 GCAATGATGATGACTGGGATACCTTTCCGGAATACCTTATTCGGCTAATCAAGAAAG 1359
DB 6424 GCAATGACGAGTATTTGGGAGAGTTTCCCGAAAGCAATTAGATCACTAATGAGTTG 6483
OY 1360 TTACCATATGTCACAAAGCTGAAAGGCGCAAGCAGACATTTGTTGGAAATCTC 1419
DB 6484 TTACCGCATGCTGGCCAGACTGAAAGGCGCTAAGGCGCGCACCTGTTCCGAAAGACGC 6543
OY 1420 ATAACTAAAACCGTTGACAGAGATACCAATGATCAATTCGATGATCTAAGAAAGAG 1479
DB 6544 ATAAATTTGGTCCCATTTGCAAGAGATGCTATGATGATTTGCTCATGACATGAAAGAG 6603
OY 1480 ATGTCAAAAGTATCTCCCGACGAAACATATACAGAGAGGCGCTAAGGTGACGATTAATC 1539
DB 6604 ACGTAAAGATTACCTGGCAGAAACACAGAAAGAAAGCCGAAAGTACAGTGAATAC 6663
OY 1540 AGGCTGAGATTCCTTGTCTACCTTACCTTTGGGGATCCAATGGGAATTAATGCTCGTA 1599
DB 6664 AAGCGGAGAAACCTGGGACCGCTTACATATGCGGATCCACCGGGGATTAATGCGCA 6723
OY 1600 GACTAATGCGGCTGCTTCCGCAAAATATCATATCTCTTTCGACATGTCAAGGGAAGAT 1659
DB 6724 GGGTTACAGCCGTTTGTCTACCCCAATTTCAACGCTCTTTTGAATGTTCCGAAAGACT 6783
OY 1660 TTGATGCAATTAATGCTGACATTTTCAACAGCGGACCCAGTATTTGAAACGACATCG 1719
DB 6784 TTGATGCAATCATACAGAACTTCAAGCAAGTGAACCGGTTACTGGAGACGATATCG 6843
OY 1720 CGTCTTTGATTAAGCGAAGCGAAGCGATATTCGCAATTTCCGCTTATGATTCCTTGAG 1779
DB 6844 CCTGTTGCAAAAAGCCAAAGCGAAGCGATATGCTTAAACCGGCTGATGATCTTGGAAG 6903
OY 1780 ACTTAGTGTGACCAACGCGCTCTTAGATTTGATAGAGCGGCGTTCCGCAATATCAAT 1839
DB 6904 ACTGGGTGTGACCAACCACTATGCTGATGATGAGTCCGCTTTGGAGAAATATCT 6963
OY 1840 CTGTGACACTTACCAAGAAAGAGTTTAAATTTGTGTCATGATGAAATCCGATATCT 1899
DB 6964 CAACCATCTGCGCCAGGTTACCGTTTCAATTTGGGGCGATGATGAAATCCGAAATCT 7023
OY 1900 TCTTAAGCTGTTTGTCAACACATGATCATGATTTGCTGACAGATTAATCACTGCTG 1959
DB 7024 TCTTCACCTCTTTTGTCAACAGATTTGAAATGCTTATCCGACAGAGATTTGAGAG 7083
OY 1960 AACGGTTAACCAAGCTCAGGTGCGGCGCTATGCGGACGATTAACATAGTCATGCTG 2019
DB 7084 AGCGCTTAAACGTCAAATGTGACATTTATGCGGACGACAACTTATATACCGAGAG 7143
OY 2020 TCGTCTCGACACTTGTATGCGGAGAGATGCGCACTTGTGCTGAACATGGAATGAAAA 2079
DB 7144 TAGATCTGACAAAAGAAATGCTGAGAGGTGACCACTGCTCAACATGAGAGTTAAGA 7203
OY 2080 TTATGATGCAATTAATTTGATATCAAAAGCACTTCTGTGGGGATTTATCTGTGTG 2139
DB 7204 TCATTTGACGAGTATCGCGCAGAGACCACTTACTTCTGCGGTGATTCATCTTGCAAG 7263

OY 2140 ACCAGATTAACAGGACACAGCTGACAGATGCGACACCTCTTAAAAAGGCTTTTAAAGCTTG 2199
DB 7264 ATTGGTTAATCTCCACAGGTGTCTGCGGCGGACCCCTTGAAGAGCTTTAAGTTGG 7323
OY 2200 GAAAACCATTTGCGAGTGTATGATTAACAGACTGCGACCGCGCGGACCTGATGATG 2259
DB 7324 GTAAACCTCTCCACCCACGATGAGCAAGACGAAAGACGAAAGACGCGCTGTGATGATG 7383
OY 2260 AAGCAATCGATGGAACGAAATTTGAATTAAGAGAGATTAATGAAAGCCGTGAAATCCA 2319
DB 7384 AAACAAAGCGGTGTTTGAAGTATACAGACACATTTACAGTGGCCGTGGCAATTC 7443
OY 2320 GATACGAGATCATATGAGAGGCTGATCATGCTCTCTGTCCAAGTTAGCCGAAAGCG 2379
DB 7444 GGTATGAGTATGACAAATCAACACCTGTCTCTGCTGATTAAGAACTTTGCGCCAGAGA 7503
OY 2380 TTAAAGCTTCAAGAGCATTAAGAGGAGCCCAATCAACCTCTAGCGCTGACCTAAATAG 2439
DB 7504 AAAGAGCAATTTCAAGCATCAGAGGGAATTAAGCATCTTACGCTGTCTTAAATAGT 7563
OY 2440 TGACGTATTAAG-----CAGGACCTAACCAACCGGAGAAATGTTTCCA 2482
DB 7564 CAGCATGATGATCTTTCATCTGATTAATACCAACACCACTGATTAAGAGATTTCT 7623
OY 2483 TACCTCAGCTGAATCTTTCACACAGTTTAACTCTAACAATTCGATGAGCTTAACGAGATCCA 2542
DB 7624 TTAATATGCTGCGCGCGCGCGCTTCCAGCCCACTGCAATGTGAGAGCGCGGAGAA 7683
OY 2543 AACCTCTAGCGCGCGCTGAGAGCGTTTCGGCCCGCTGCTGCTTAAATGGAAGAT 2602
DB 7684 GAGGACGAGCGCGCGCGATGCTGCGCAATATGAGCTGCTTCCCAATTCAGCAATCTGA 7743
OY 2603 CTTAAGAGTGCATATGATCACTTCAATCAACAAAC--GATCACTTAATCCGCGCCAG 2660
DB 7744 CCACAGCGGTGATGCTTATGATTTGACAGGCAATAGACTTAACCCACCCAC 7803
OY 2661 GTCCACCGCCAAAGAAAGAGTGTCTCTTAAGCCAAACCTTACTGACCTTAAAGAA 2720
DB 7804 GCCCGCGCGCGCGCGAGAGAGAGCGCGCCAAAGCAACCTGAAAGGAAACCA 7863
OY 2721 AGAAGCAGCAAGCCAAAGAGAGCAAGAAC--GCAGCTTAAACAGAGAAACGACACGTA 2777
DB 7864 AAACACAGAGAAAGAAAGAAAGCAACCTGCAAAACCCAAACCCGAAAGAGACAGCGTA 7923
OY 2778 TGTGATGAAGTGTGAGTGTGACAAAGACATTTCCATC---ATGCTGAACGCGCAAGTCA 2834
DB 7924 TGGCATTAAGTTGAGGCGCAACAGCTTTCAGCTGCAAAATAGAGACGAGATGTCA 7983
OY 2835 ATGATATGCTGCGTGTGCGAGAGAGCTGATGAACCACTTCACGTTGAAGAAAA 2894
DB 7984 TCGGACGACACTGCGCATGGAAGAAAGTATTAACCACTTCACTGTAAGAGAACTTA 8043
OY 2895 TTGATTAATGACAAATTAGCGCGCGTGAATTTGAAGAGCTAGCATTAACGATTTGAGAT 2954
DB 8044 TTGACCAACCTGTGTATCAAAAGCTCAAAATTCACCAAGTGTGACATATACGATGAGAT 8103
OY 2955 ACGGACGCTTCCCAAGAACTGAATCAAGACGCTGCAATGACACGACGCAAAACCAAC 3014
DB 8104 TCGCACAGTGTCCGTTCAACATGAGAGTGAAGGCTTACCTTACCACTACAGTGAAGCCCTG 8163
OY 3015 CGGCGCTTACAACTGGAGACCAAGCGCGAGTCAATATGAGAAAGGAGATTACCGTAC 3074
DB 8164 AAGGATTTTACAACTGGGACCAAGAGCGGTGTGATATAGTGAAGAGAGATTTTACATTC 8223
OY 3075 CGAGAGAGTGGCGGGAAGGCGACAGCGGAACCGATCTCTGACAAACAGAGCAGAG 3134
DB 8224 CCGCGGAGTATGAGAGGACAGAGAGCAATGTGTGCTGATTAATGATTAATCAAGCCGAG 8283
OY 3135 TTGTGCTAATTTGTTTATGAGAGTGTCAATTAAGAGGACGCGCTTCAAGTGTCA 3194
DB 8284 TTGTGCGATATGCTCTCGAGAGGCTGATGAGGGAACAAAGACCGCTTTCGCTGCTCA 8343
OY 3195 CTTGGAACCAAGAAAGGCTGACATTAAGGATACCCCGAAGGTTCTGAACCGTGT--- 3251


```
Db 8244 CTTGGATGAGAAAGGGAAGCAATCAAGAACACCCCGAAGGGAAGAGAAAGTGTCTG 8403
Qy 3352 -----CACTGATACAGCCGCTATGCGTCTTGAATGTCACTGTTCCCATGCCAACAAC 3305
Db 8404 CTGCACCACTGTGTACCGGCCATGTGTCTTGGAAAAGTGAAGCTTCCCATGCATCGCC 8463
Qy 3306 CACCGTGTGTATTTCACTGACGCGAAGAACGAACTGCACTGCTGCAAGAGAAGTGTG 3365
Db 8464 CGCCCACTGTGTACACCGCGCAACCTATCAGAGCTCTTGACATCTCTGAAAGAAAGTGA 8523
Qy 3366 ACAATCCAAATTACGACACGCTGTGAGAAAGTCTTGAAAT---GTCCATCAGCCGCGC 3422
Db 8524 ACCACGAGGCTTACGACACCTGTCTCAACGCCATATGCGGTGGGAGTGTCCGCGAGAA 8583
Qy 3423 CCAAAACGAAGCATTAACGATACCTTACATGACCAAGTCTTACTGCGGTTCTGCGCT 3482
Db 8584 GTAAAAGAAAGGCTACGTACACACTTACCTTGACAGCCCTGACTTGGGCAATGTCTGT 8643
Qy 3483 ATTGCAGACTCAACGCGGTGTTCAGCCCAATAAAATTGAAAGCTGTGGGACGAAT 3542
Db 8644 ACTTTCACATCTGAAACCGTGTCTTACCCGATTAAATGACAGAGGTCTGGGATGAAAG 8703
Qy 3543 CTGATGATGATCGAATTAGAAATCGAGGTCTGCGCAATTCGCTACATCAGGCAAGCA 3602
Db 8704 CGAGCAACAACACATACGATACAGACTTCGCCCAAGTTTGGATACGACCAAGCGGAG 8763
Qy 3603 CTGCGGATGTACCAAAATCCGTTACATGTCTTTCGACACGACCATGACATCAAGAAAG 3662
Db 8764 CAGAGACTCAAAATAGATACCGCTACATGTGCTCGACAGAACTATCTGTCAAAAGAG 8823
Qy 3663 ACAGATATGAGAAAATAGCTATCAGACATCTGAGACCTGCGGTCGTCTTGGCCCAAG 3722
Db 8824 GCACCATATGATGATCAATCAAGATCAGACCTTACGACCGTGTAGAAAGCTTATAGCTAACAAAG 8883
Qy 3723 GGTACTTCTGTGTAGCTCAATGTCTCTCCAGGTGACAGTGAACCGTCAATCAAGAGCG 3782
Db 8884 GATCTTCTCTCGCGCAAGTGTCTCTCAGGGGACAGGTAAAGCTTATGATACAGAGTA 8943
Qy 3783 GAGCATCTGAGAAATTCATGACACCGTGAGAAAAGATGAGAGAAAGTTGTCCGTAGAG 3842
Db 8944 GCAACTCAGCAACGTCATGCAATGAGCCGCAAGATTAACCAAAATTCCTGTGGACCGG 9003
Qy 3843 AGAGATCTTGTGCCACCGCTCATGGAAGCTGTAAAGTGCACGCTTATACATCACT 3902
Db 9004 AAAAATATGACTTACCTCTCCGTTACCGGTAAAGATTCCTTGCACAGTGTACACCGTC 9063
Qy 3903 TGAAGAGACGTCTGCGCGGTACATAACCATGACACAGGCCAGGCCACACCGCTATAGT 3962
Db 9064 TGAAGAAACAAACCGCGGTACATCACTATGACACAGGCCGGAACCGCATATACAT 9123
Qy 3963 CCTATCTGAGAGAACGCTCAGGCCAAGTGTACATTAACCACTTCTGGCAAGAACGTCA 4022
Db 9124 CCTATCTGAGAGAAATCATCAGGAAAGTTTACGCGAAACCAATCCCGGAAAGAACATTA 9183
Qy 4023 CCTACGATGTAGTGTGCGACATACAGACAGGATGTGAGACCGCGGAAGAAAGATGA 4082
Db 9184 GGTACAGTGTAGTGTGCGCGCATTTACAGACCGGAACCGTTACGACCCGTTACCGAAATCA 9243
Qy 4083 ACAGCTGCACTTAAAGCAAAACAGTGCATTTGCTTACAGAGGACCAACGAATGGGTCT 4142
Db 9244 CGGCTGACCGCATCAAGCAGTGTGTGCTTAAAGAGCAACAAACGAAGTGGGTCT 9303
Qy 4143 TCAACTGCGCGGATTTATTTGGCACACAGACCACTCAGTGCAGAGTAAATTTGCACATTC 4202
Db 9304 TCAACTGCGCGGACTCTGATCAGACACGCGCAACACCGGCCCAAGGAAATTTGATTTGC 9363
Qy 4203 CATTCGCTTGACACCGACAGTGTGCGCGGTTCGTTAGCTCAACAGCTACAGTCAAGCA 4262
Db 9364 CTTTCAAGCTATCCGAGTACTGACATGTGTCTGTGCGCAAGCGCGCAACGATGATAC 9423
Qy 4263 AGTGTTCAAAGGATCAACCTTCAACCTGACATGCAACCAATTTGCTGACAAAGCA 4322
Db 9424 ACCGCTTAAACACATCAGGCTCCCAATTAGACACAGACCAATCTGACATTTGCTCACCA 9483
Qy 4323 GAAATTTGGGGGTGTGAGAGAGCAACAGAGAAATGATTTACAGGCTTACATCCAGA 4382
Db 9484 GGAATCTAGGGGCAAAACCGGAACCAACACTGATATGATATCGAAACAGGTTAGAA 9543
Qy 4383 ATTTTCTGTGGGCGAAGAGGCTGAGTACGATGAGGGGTAAACATGAACCACTCAGAG 4442
Db 9544 ACTTCAACGCTGACCGAATGTGCTGATATCAATATGGGGGAATCAAGAACCAATGAAGG 9603
Qy 4443 TGTGGGCCAGAGAGTGTGCAACAGGCAACCAATGATGAGCGCGCATGATCATATCC 4502
Db 9604 TCTATGCCAAAGTCTGCAACAGAGACCTTACGATGAGCCCAAGAAATATGATACAGC 9663
Qy 4503 ACTATATCAATGCGCATCCAGTGTACATGTCATGTGTGTGTGTGTGTGTGTGTGTGTGT 4562
Db 9664 ATTAATCATCTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9723
Qy 4563 TCTGTAGGCACTGATCATATCAGACAGCTTGCATCGCCAAAGCAAGAAAGACTGCTGA 4622
Db 9724 TGATGATTTGGGTAACTGTGGACGATTAATGTGTGTGTAAAGGCGCGGTGAGTGTGCA 9783
Qy 4623 CGCCATACGCGCTTGACACCGAAGCAAGGTAACCAACATTAAGCGGTTTGTGTGTGCA 4682
Db 9784 CGCCATATGCTGTGCGCCCAAGTCCGTGATTCCACTTGTGCTGCACTTGTGTGTGTGTG 9843
Qy 4683 TTGGGCCAACCAACGCTGAAACATTTGAGAACTTGAACCATCTGTGGTTTAAACACC 4742
Db 9844 TTAGTGTGCTGATGTGAAACATTCACGAGACATGATGATTTATTTGTGTGTGTGTGTGT 9903
Qy 4743 AACGTTTCTGTGGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4802
Db 9904 AGCGTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9963
Qy 4803 GCTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4862
Db 9964 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10023
Qy 4863 CTTTGAACATGCGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4922
Db 10024 CTTAGCAACATGCGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10083
Qy 4923 AACCGAGGTTTACGCGCACTTAACTGAGATCAGGTCGTCTACATCGGAATTAACAC 4982
Db 10084 AAAGGAGGAGTACCGCCGCTCAATTTGGAGATTAATGTATGTCTCTCGAGGTTTGTG 10143
Qy 4983 CTTCACTAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5042
Db 10144 CTTTCAACCAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10203
Qy 5043 TTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5102
Db 10204 TCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10263
Qy 5103 TTTTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5162
Db 10264 TCTTTGAGGGGTGTATCCCTTCAATGTGGGAGAGCAATGTTTTGGACAGTGA 10323
Qy 5163 ACACACAATGATGAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5222
Db 10324 ACAGCCAGATGAGTGAAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10283
Qy 5223 TCGACCTAAAGTTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5282
Db 10384 AGGCGATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10443
Qy 5283 CCACCGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5342
Db 10444 CTAACAGTTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10503
Qy 5343 AGGTATAGCAGGGCGCATATCAGCGCTTTTTCACCTTTGACCATTAAGGTGTGTGTGT 5402
Db 10504 AAGTCATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10563
```


QY	5403	GAAGGGGCTGTGTTTAACTACAGACTTCCCTGATATGAGCTATGAACCGAGAGCGT	5462
Db	10564	ATCGGGCCCTGGTGTCAACTATGACTTTCCGAAATACGAGGAGTGAATCCAGAGCGCT	10623
QY	5463	TCGGCGATTAATTCAGACATCTCCGCTTGATGCTACAGACATATGAGCCGCACTGACATAC	5522
Db	10524	TTGGAGACATTTCAAGACTACCTCCTTGACTAGCAAAAGCCTCATCGCAGAGACATTA	10683
QY	5523	GGCTGCTGAAGCCTTCTGTCAAGAACATCCAGTCCCTTACACCAAGCATACAGGT	5582
Db	10684	GGCTACTCAAGCCCTTCGCGCAAGAACGTGATGCTCCGTACACGACGGCCGCACTCGGAT	10743
QY	5583	ATGAAATGTGAAAGAACATCAGACGACCCCTGCAGAAAGAACAGACCATTTGATGTA	5642
Db	10744	TCGAGATGTGAAAGAAACATCAGGCGCGCCACTGCAGAAAGCCCTTTTGGGTGTA	10803
QY	5643	AAATTGAAGTGAAGCCTCTGCGAGCGCTCTAATCTGTGTTACGGGCAATCCCTATCTCGA	5702
Db	10804	AGATTGCAATGTCATTCGGCTTCGAGCGGTGGACTGTCTATACGGGAATTTCCATTTCTTA	10863
QY	5703	TTGACATCCCTGATGACAGCTTTTGTGATCATCAGAAATCACCACAAATTTTGAAGTTA	5762
Db	10864	TTGACATCCCGAAGCGCTGCTTATACAGACATCAGATGACACACGCTCTCAACAGTGA	10923
QY	5763	GCTGCACAGTACAGACATGCAATTTATCTGACAGCTTTGGTGTCTTAACATTTTACAT	5822
Db	10924	AATTGATGTCACTGATGATGCACTTATTCACGGAATTCGAGGAGATGGCTACCCTGCAAT	10983
QY	5823	ACAAAGCTGACAGGAGGAGACATTTGTCAATTCATCTCCACTCACAGACAGCTGTTTGA	5882
Db	10984	ATGATCTCGAACCGGAAGACAAATGCCCTGTACATTTGACTTGACACACGCAACCTTCC	11043
QY	5883	AGGAAGGACCAACATGTGACTGCGGTAGGACAGCAATACACTACTTTTACGACATGCA	5942
Db	11044	AAGAGTGCAGACATTCATGCTCTGGAAGAAAGACGGGTGACATGACACTTCAGACCGCGA	11103
QY	5943	GCCCAACAAGAAATTTATATGTTTCCGTATGGGGCAAGAGTCCACTGCATATGCTGAT	6002
Db	11104	GCCCAAGCGGAACCTTATGTGATCGCTGTGTGTGAAGAACAAACATGCAATGACGAAT	11163
QY	6003	GTAAACCAACCGCGCGACACATATAATGGAAGAACACATAAAGTCGACCAAGAAATTCAGG	6062
Db	11164	GCAAAACCAACGCTGATCATATGTGTAGACACCCCGCACAAATAAGACAAAGATTTCCAG	11223
QY	6063	CGGCAGTTTCAAAACATCTTGAACGTGCTGTGCACTGTTTGGGGAGCATCAATCC	6122
Db	11224	CCGCACTCATAAAACCTTCATGAGAGTGCTGTGTTGGCCCTTTTCGGCGGGCCTCGTCC	11283
QY	6123	TCATGTGTTGAGACTTATATGTGTGCTGTGACGCTATATGCTTATAAACACACGTAGAT	6182
Db	11284	TATTAATTTATAGCACTTATATATATTTTGTCTTGACGATGATGCTGATACACACGAAAGT	11343
QY	6183	GACTGACCGCGGACACTGA-CATAGCGGTAAACCTGATGATCTTCCGAGGAAGCGTGT	6241
Db	11344	GACCGCTACGCCCCCAAGACACCGACGCAAAACTGATGTACTTCCGAGAACTGATGT	11403
QY	6242	GCATATGCCACGGCGCGCT-----TGAACT	6268
Db	11404	GCATTAATGCAATCAGGCTGTGTATATTAATGATCCCGCTTACCGCGGGCAATATAGCAACACC	11463
QY	6269	AAAACTGATGATTTTCCGAGAGACACATGCTAATAATGCTGTGCACTGT	6318
Db	11464	AAAACTGACGTATTTTCGAGGAAGCGCACTGTGATATATGCTGTGGCACTGT	11513

```

RESULT 9
ABX81526
ID      ABX81526 standard; cDNA; 11740 BP
XX
AC      ABX81526;
XX
DT      23-APR-2003 (first entry)

```

XX	Sindbis virus cDNA #3.
DE	
XX	Eukaryotic layered vector initiation system; cytopathic effect; gene; ss;
KW	alphaviral nonstructural protein 2; RNA vector replicon; alphavirus RNA;
KM	macromolecular synthesis; alphavirus RNA polymerase recognition sequence;
KW	3' polyadenylate tract; cellular macromolecular synthesis; cell death;
KX	Sindbis virus.
OS	Sindbis virus.
XX	
PN	US6465634-B1.
XX	
PD	15-OCT-2002.
XX	
PF	08-OCT-1999; 99US-0415900.
XX	
PR	06-OCT-1997; 97US-0944645.
PR	05-APR-1996; 96US-0628594.
PR	24-JUN-1996; 96US-0668953.
PR	12-JUL-1996; 96US-0679640.
PR	04-APR-1997; 97US-0833148.
XX	(CHIR) CHIRON CORP.
PA	(UNIV) UNIV WASHINGTON.
XX	Dubensky TW, Peio JM, Belli BA, Schlesinger S, Dryga SA, Frolov I;
XX	WPI; 2003-147073/14.
DR	
PT	Eukaryotic layered vector initiation system, for gene therapy, has
FT	alphaviral nonstructural protein gene having mutant nonstructural
PT	protein 2 gene, which reduces host-cell directed macromolecular
PT	synthesis -
PS	
XS	Example 1; Fig 8A-8E; 161pp; English.
XX	
CC	The invention relates to a eukaryotic layered vector initiation system,
CC	comprising a nucleic acid sequence encoding all four alphaviral
CC	nonstructural proteins and including an altered sequence encoding for
CC	nonstructural protein 2, such that when the altered sequence is operably
CC	incorporated into an RNA vector replicon, the time required to reach 50%
CC	inhibition of cellular macromolecular synthesis in cells is increased, as
CC	compared to an RNA vector replicon having a wild-type alphavirus
CC	nonstructural protein 2. The initiation system comprises a 5' promoter
CC	which directs synthesis of alphavirus RNA in vivo from cDNA, a 5'
CC	sequence which directs transcription of alphavirus RNA, a nucleic acid
CC	sequence which operably encodes all four alphaviral nonstructural
CC	proteins, an alphavirus RNA polymerase recognition sequence and a 3'
CC	polyadenylate tract. The eukaryotic layered vector initiation system is
CC	useful for stimulating an immune response within a vertebrate, for
CC	protein expression and gene therapy. The system exhibits reduced, delayed
CC	or no inhibition of cellular macromolecular synthesis, thus permitting
CC	its use for protein expression and gene therapy with reduced, delayed or
CC	no development of cytopathic effects or cell death. This sequence
CC	represents Sindbis virus cDNA, used in the scope of the invention.
XX	
SO	Sequence 11740 BP; 3345 A; 3049 C; 2908 G; 2438 T; 0 other;
Query Match	36.7%; Score 2380.4; DB 25; Length 11740;
Best Local Similarity	64.6%; Pred. No. 0;
Matches 3693; Conservative	0; Mismatches 1986; Indels 41; Gaps 8;
Oy	580 ACTGACGGTATGGAAGCGGAGCGGTATTCTTCATCGGAAACGCGCAAGTCACTTC 639
Db	5746 ACTGACTTAACCGGGGTAGTGCGTACATTTTTCGACGCACACAGGCCCTCGGCATCTTC 5805
Oy	640 AACGAAATCAGTAGTCACTCAATGTAATACTACAGAACAATCATATTGGATCGGCGCTTCATG 699
Db	5806 AAAAGAGTCGCTTCTGCAGAACCACTTACAGAACCGAAGCTTGGAGCGCAATGCTCGG 5865
Oy	700 AGAAGTATTACGCCCGCGCGCTGCATTCGAAAAGAGAAAATGTTACAGAGAACTGC 759

Dh 5866 AAGAAATTCATGCCCCGGTGTGACACGTGCAAGAGAAACAATCAACTCAGTACC 5925
Qy 760 AATTATGCGCTCTGAAGGAATAAGACAGTATCATACGAAAAAGTAAATAATATGA 819
Db 5926 AGATGATGCCCAACGGAAGCCAAACAAAGTAGTACAGTCTCGAAAGTAAATAATCAGA 5985
Qy 820 AAGCAATTAACAGCGAGGAGCTCATTTCTGATTTGGGACACATATCATCAGAAAGTGA 879
Db 5986 AAGCATTAACACATGAGGAGTACTGTGAGAGACTACAGCTATTAACCTCTCCACAG--- 6042
Qy 880 ATCTGTGAGTGTAAACAGTCAATATCTGTACCAATCTACTCTGTCAACGGTAAATTA 939
Db 6043 ATCAGCCAGAAATGCTATTAAGATCACTATCCGAAACATGTACTCCAGTACGGTACCGG 6102
Qy 940 ACAGGTTTACATCTGAGAGAGTCCCGGTTAAAGCTGAACCTTAAGTTATCCAAAGAAAT 999
Db 6103 CGAACTACTCGAATCCACAGATTCGCTGTAAGTCTGTGTAACACTATCTGTGATAGAACT 6162
Qy 1000 ACCCTAAGTAGGACAGTATATGTATTAACAGATGAATACGATGCGTATCTGACATGGTGG 1059
Db 6163 ATCCGACAGTAGCATCTTATCAGATTAAGTACGAGTAGAGTGTCTTAATTGATATGGTAG 6222
Qy 1060 ACGGCGATCGTGTCTGTCTAGATAACGCCATTTTGTCCGGCTAAACCTGAAGACTACC 1119
Db 6223 ACGGAGCAGTGCCTGCTGATCTGCAACCTTCTGCCCCGCTAAGCTTAGAAGTTAAC 6282
Qy 1120 CAAGAAGCATAGCTATTTGAGAGCCAGATTAAGATCAGCCGTCCCATGCGCTATACAGA 1179
Db 6283 CGAAAAAACAAGATATAGAGCCCGCAATATCCGAGTGCCTTCATCAGCGATGCGA 6342
Qy 1180 ATACATTACAAATGTATTTGGCTGACAGTACTAAAGAAATGCAACCTTACCCAAATGC 1239
Db 6343 AACGCTACAAATGTGTCTATTTGCCCACTAAAGAAATTCGACCTCAGCGAGATGC 6402
Qy 1240 GAGAAATTAAGTCTGTAGATTTGGCGGAGCTTAATGTTGTTTCAAGAAATACGAT 1299
Db 6403 GTGAAGTCCCAACACTGAGACTCAGCGACATTCATATGTGAAATGCTTTCGAAATATGAT 6462
Qy 1300 GCATGATGATGACTGAGGATACCTTTGGCGAATACCTTATGGGCTAATCTACAGAAAG 1359
Db 6463 GTAAATGCGAGTATTTGGAGAGATTCGCTCGAAGCCCAATTAAGATTAACCTAAGTTTG 6522
Qy 1360 TTACGCAATATGTGACAAAGCTGAAAGGCGGAAGCAGCAGATTTGTTTSCAATATC 1419
Db 6523 TCACCGCATATGTAGCTAGACTGAAGGCCCTTAAGCCGCCGCACTATTTGCCAAAGAGCT 6582
Qy 1420 ATAAATCTAAACCGTGTGAGAGATACCAATGATCAATTTGTCATGAGATCTAAAGAG 1479
Db 6583 ATAAATTTGCTCCATTTGCAAGAGTGCCTATGATAGATTCGTATGAGCAATGAAGAG 6642
Qy 1480 ATGTCAAGTTATCTCCCGGACGAAACATACAGAGAGCGGCTTAAGTGCAGTTATTC 1539
Db 6643 ACGTGAAGTTTACACAGGACGAAACACAGAGAAAGAACCGAAAGTAAAGTGAATAC 6702
Qy 1540 AAGGTGAGATCCCTTGTACCGCTTACCTTTGCGGGATCCATCGGGATTTAGTCCGTA 1599
Db 6703 AAGCCGAGAAACCTCTGCGCACTGCTTAATGCGGATTCACCGGAAATTAAGTGCCTA 6762
Qy 1600 GACTGAATGCGGTCTTGTGCAATATCATCTCTTGTGACATGTCAAGCGGAAGAT 1659
Db 6763 GGGTTACGGCGCTGTGCTTCCAAACATTCACAGCTTTTGAATGATGTGCGGAGGAT 6822
Qy 1660 TTGATGCGATTAATGCTGAACATTTCCACACGCGGACCCAGTATTTGAAAAGGACATCG 1719
Db 6823 TTGATGCAATCATAGCAGAACATTCACAGAGCGACCCGCTACTGGAAGCGATATCG 6882
Qy 1720 CGTGTGTTGAATAAGGAGAGACGATCGCATTTGCGCGTGAATGATCCCTGAGG 1779
Db 6883 CATATTCGACAAAGCCAGAACACGCTATGCGCTTAACCGGTCTGATGATCTTTGAGG 6942
Qy 1780 ACTTATGCTGACCAACCGCTCTTAATTTGATAGAGCGCGCTTGGCAATATCAT 1839
Db 6943 ACCTGGGTGTGATCAACCACTACTCGACTTATGATGAGTGGCCTTTGGAGAAATATCAT 7002

Qy 1840 CTGTCACCTTACCTACAGAAACGAGTTTAAATTTGGTGCATGATGAATCCGGTATGT 1899
Db 7003 CCACCCATCTTACTACCGGATCTCTGTTTAAATTTGGGGGAGATGAATATCCGAATGT 7062
Qy 1900 TCTTAACTGTTTGTGCAACAACACTACTCAATATCATATGATTTGTAGCAGATCTACGTG 1959
Db 7063 TCTTCACACTTTTGTCAACACAGTTTGAATGTGTTATTCGACAGAGTACTTAAG 7122
Qy 1960 AACGTTTAAACAGTTCAGCGTGGCGGCTCTATCGCGACAGTAAATATAGTGAATGTG 2019
Db 7123 AGCGGCTTAAACGTCAGATGTGCAAGCTTCATTTGCGACAGAACATATACATATGAG 7182
Qy 2020 TCGTCTCGACACCTTGATGCGGAGAGATGCGGCCACTTGGCTGAAATGATGAATAAAA 2079
Db 7183 TAGTATCTGACAAAGAAATGCTGAGAGGTGCGCCACTGCTCAATGAGAGTTAAGA 7242
Qy 2080 TTATTTAGTAGATTAATTTGTATCAAAACACCTACTTCTGTGGGGAATTAATCCTGTGG 2139
Db 7243 TCATGACGAGTATCGGTGAGAGACCACTTAATCTTGGGGGATTAATCTTGGCAG 7302
Qy 2140 ACCAGATTAACAGGACACGCTGCAAGATCGCAACCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7303 ATTCGTTACTTCCACAGCGTGGCGGCTGGCGGATCCCTGAAAGGCTGTTAAAGTTGG 7362
Qy 2200 GAAACCAATGGCAGTGAATGATACCCAGACTGCGGACCGCGCGGCACTGATGATG 2259
Db 7363 GTTAACCGCTCCACGCCAGACGAGCAAGACAGAACAGACAGAGCGGCTCTGCTAGATG 7422
Qy 2260 AAGCAATGCAATGGAACAGAAATGGAATTAACGACGATAGTGAAGGCGGTAGAAATCA 2319
Db 7423 AAACAAAGGGGTGTTAAGATAGTATTAACGACCTTAAGCAGTGGCGTGAACGCC 7482
Qy 2320 GATACGATATCATCTGACAGGCTGATCATCAGTCTCTGTCCACTTACCGGAAAGCG 2379
Db 7483 GGTATGAGGTAGACAAATATTAACCTGTCTTACTGATGAGAACTTTTGGCCAGAGCA 7542
Qy 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGGCCCAATCACCTCTACGGCTGACCTTAATAG 2439
Db 7543 AAMAGCAATTCAGAGCATCAGAGGGAATTAAGCATCTCTACGATGTGCTCTTAATAGT 7602
Qy 2440 TGAAGTATGACACGCACTAC-----CCACCGGAGAAATGTTTCCATACC 2486
Db 7603 CAGCATAGTACATTTATCTGATCTAATATACAAACACACACATGAATAGAGATTTCT 7662
Qy 2487 CTCAGCTGAACCTTTCACACGATTTTACCTTCAATCCGATGCTTACCGAGATCAAAAC 2546
Db 7663 TTAAATGCTTCGCGCGCGCGCCCTTCCCGGCCCCCACTGCAATGTGAGGCCCGGAGAA 7722
Qy 2547 CTCTAGGCGCGCTGAGGCGGTTTGGGCCCGCGCTGGCTCAATGGAATCTTTA 2606
Db 7723 GAGAGCAGGGGCGCCGATGCTGCGCAAGGGCTGCGCTCTCAATTCAGBAATCTGA 7782
Qy 2607 GAGAGTGAATGATCACTTGAATCTTCAACAAACGATCACTTAATCCGCGCAGTGTCA- 2665
Db 7783 CCACAGCGTCAAGTGCCTGATGATGACAGGCAATAGACCTCAACCCCAAGTCCAC 7842
Qy 2666 -----CGCCAAAGAAAGAAAGTCTCTTAAGCCAAACCTTCACTCAAGCTTAAAGAA 2720
Db 7843 GCCCGCACCGCGCCAAAGAAAGGAGCGGCCCAAGCAACACCGAACCGAAAGAAACAA 7902
Qy 2721 AGAAGCAGCAAGCCAAAGAGACGAAAC---GCAAGCTTAAACAGGAAACGACAACTGA 2777
Db 7903 AAGAGCAGAGAAAGAAAGAAAGACACTTCAAAAACCGGAAAGAGACAGCGCA 7962
Qy 2778 TGTGTATGAAGTTGAGTGCAGACAGACATTTCCGATCATG---CTGAACGCGCCAACTGA 2834
Db 7963 TGGCACTTAATGTTGAGGCGGACAGATTTGTTCAAGCTCAAGAAACGAGAGAGATGTCA 8022
Qy 2835 ATGATATGCTGCTGTGTTGTGAGAGAAAGCTGATGAACCACTTCACGTTGAAGAAATA 2894
Db 8023 TCGGGACGACACTGGCCATGGAAGAAAGTAAATGAACCTTGCACGTAAGAAAGAACCA 8082

QY	2895	TTGATTAATGACCAATTAGCGCCGCTGAATTTGAGAAAGCTACGATGACACTTGGAGT	2594
Db	8083	TCGACCAACCTCTGCTATCAAAAGCTCAAAATTTTACCAAGTCCTGACATACGACATGAGT	8142
QY	2955	ACGCGACAGCTTCCCCAGAACATGGAATTCAGACACGCTGCAGTACACGACGACCAAAACAC	3014
Db	8143	TCGCAAGATTGCCAGTCAACATGAGAAAGTAGGCAATTCACCTAACCAAGTGAACACCCCG	8202
QY	3015	CGGCGCTTTCACACTGGCACCGGCGCAGTCCAGTATGAGAAATGGAGATTTACCCGTAC	3074
Db	8203	AAGGATTTCTATACTGGCACACCGAGCGGTGCGATATAGTGAGGTATTAATTCATCC	8262
QY	3075	CGAAGAGAGTGGCGGGAAAGGCCACACGCGAABACCGATCTCTGACAAACAGAGCGCAG	3134
Db	8263	CTCCCGGAGTAGGAGCGCAGAGAAACACCGGATCTCCGATATGATTAATCTCCCGTGGG	8322
QY	3135	TTTGCGCTATGTTCTTAGAGAGTGCAAATGAGGGGACGCGATACGGCGCTTTCACTGATCA	3194
Db	8323	TTTGTCGCGATAGTCTCGGTGGCGCTGATGAGGAACACGAATCTCCCTTTCCGTGCTCA	8382
QY	3195	CTTGGAAACGAGAAAGGGGTGACCAATTAGGATACCCCGAGAGTTCTGAAACCGGTGT---	3251
Db	8383	CCTGGAAATAGTAAGGGAAGACAAATTAAAGACACCCGGAAGGGAACAGAAAGTGTCCG	8442
QY	3252	-----CACTAGTTACACGCGCTATGCGTCTTTGGAATGTCAAGTCCCATGCAAC	3305
Db	8443	CAGCAACCACTGGTACCGCAATGTGTTGTCTGGAAATGTGAGCTTCCCATGCAACGCGC	8502
QY	3306	CACCCGTGTGCTATTACCTGACGGCAGAACGAACATGCACGTCTCGAAGAGAACGTCG	3365
Db	8503	CGCCCACTGCTTAATCCGCGCAACTTTCAGAGCCCTCGACATCTTTGAAGAGAACGTGA	8562
QY	3366	ACAAATCCAAATTTACGACACGCTGTGAGAACGTCTTGAATGTCCATCAAC--GCCGCG	3422
Db	8563	ACCATGAGGCTTACAGTATCCCTGTCAATGECATATTTGCGGTGGGATCTGTGCGAGAA	8622
QY	3423	CCAAACGAGATTACCGATACCTTACATGACCAAGTCCCTACTCTGGGGTTGCCCGT	3482
Db	8623	GCAAAAGAGGCTATTGACACATTTTACCTCGACAGGCCCTTACTTGGGCAACATGCTGT	8682
QY	3483	ATTGCAGACACTCAACCGCGTGTTCAGGCCCAATTAATAATTGAAACGTGTGGAGCAAT	3542
Db	8683	ACTGCACCACTACTGTACCGGTCTTACGCCCTGTTAAGATGACACAGTCTGGGACGAAG	8742
QY	3543	CTGATGATGATTCGATTAGAAATCCAGGCTCTGGCACAATTCGGCTAACATTCAGGACGCA	3602
Db	8743	CGAGCGATTAACACATACGCAATACAGACTTCCGCCACTTTGGATTAAGACCAAGCGGAG	8802
QY	3603	CTGCGGATGTCACCAAAATTCGGTTACATGTCTTTGCACAGACGACCATGACATCAAGGAG	3662
Db	8803	CAGCAAGGCGCAAAACAAAGTACCGTACATGTGCTTAAGCAGAGTACACCGTTAAAGAG	8862
QY	3663	ACAGTATGGAATAAATAGCTATCAGGCACATCTGGACCTCGCGCTCTTGGCCACAAG	3722
Db	8863	GCACCATGAGATCAAGATTAAGTACACTTGAAGCCGTGTGAAGGCTTAGCTAACAAAG	8922
QY	3723	GGTACTTCTCTTAGCTCAATGTCTCTCAGGCTGACAGTGAACCGTCAAGTATCAAGCGC	3782
Db	8923	GATACTTTCTCTCGCAAAATGCCCTCCAGGGGACAGCGTAAACGGTTAAGCATAGTAGTA	8982
QY	3783	GAGCATCTGAAATTCATGACCGGTGAGAAAAAGATCAGAGGAAAGTTGTCTCGTAG	3842
Db	8983	GCAACTCGACAAACGTACGTACACTGGCCGCAAGATTAACAACAAATTTGTGGGAAGGG	9042
QY	3843	AGGAGTACTTGTTCACACCCGTCCATGGAAGCTGTGAAGTGGCAAGTTTAGATCACT	3902
Db	9043	AAAAATATGATCTACTCTCCGTTACCGTTAAAAAATTCCTTTCGACAGTATACACCGTC	9102
QY	3903	TGAAGAGAGCTCTGCGCGGTACATAACATATGACAGGCGCAGGCGCCACACGCGTAATAGT	3962
Db	9103	TGAAAGAAACAATGCAAGGTACATCACTATATGACAGGCGGAGACCGCACGCTTATATCAT	9162
QY	3963	CTTATCTGAGAAAGCGTCAAGCGAAGTGTACATTAACAACCTTCTTGGACGAAGCTCA	4022

Db	9163	CTTACCTGAAAGAAATCATCAGGAAAGTTTACGAAAGCCGCCATCTGGGAAACAATTGA	9222
OY	4023	CCTACGAATGTAAAGTGTGGCGACTACAGACACAGGTATCGTGAACGCGCAACGAAGTGA	4082
Db	9223	CGTAAAGATGCAAGTGTGGCGCACTACMAAGCCGGAAACGTTTGCACCCCGACCGAAATCA	9282
OY	4083	ACGCGTGACTAAAGCAAAACAGTGCATTGCTCTACAGAGCGCAACAAAGAAATGGGTCT	4142
Db	9283	CTGGTTGACCGCCCATCAAGCAGTGGTGGCTTATTAAGAGCGCAAAAGAAATGGGTCT	9342
OY	4143	TCAACTCGCCCGGATCTTATTTAGGCACACAGACCCTCAGTGCAGGTAATTGCAATTTC	4202
Db	9343	TCMACTCACCGGACTTGATTCAGACATGACGACCAACCGCCAAAGGAAATTCATTATTC	9402
OY	4203	CATTCCGCTTTGACCCGACAGTCTGCCCGGATTCGGTTACCTACACAGCCCTACAGTACGA	4262
Db	9403	CTTTCAAGTTGATCCCGAGTACTCTGATGATGCTCCCTGTTGCCAGCGCCGAATGTAAATAC	9462
OY	4263	AGTGGTTCAAAAGGCATCAACCTCTCCACCTGACTGCAATGCGAACCAATTGCTGACACGA	4322
Db	9463	ATGGCTTTAAACATCAAGCCTCTCCATTATGATATACAGACCATTGACATTGGCTCACCA	9522
OY	4323	GAAATTTGGGGCTGGGAGACGACGCAACAGCGAATGATTTACAGGCTTACATCCAGGA	4382
Db	9523	GGAGCTAGGGGCAAAACCCGGAACCAACCACTGAATGATGCTCGGAAAGAACGCTAGAA	9582
OY	4383	ATTTTCTGTGGGGGAGAGAGGGGCTGGAGTACGTATGGGGTAAACATGAACAGTACAG	4442
Db	9583	ACTTACCGGTGACCGAGATGGCTCTGGATATACATTTGGGAAATCATGAGCCAGTACGG	9642
OY	4443	TCTGGGCCCAAGAGTCTGGGACCCAGGCGAACCAATGATGAGCGCATGAGATCATCTCC	4502
Db	9643	TCTATGCCCAAGATCAGACACGAGAACCTCTCAGGATGGCCACGAATATGTAACGC	9702
OY	4503	ACTATATCATTCGGCATCATCAGTCTTACACTGTATGCTGTGTGTGCTGCTCTTGGCTA	4562
Db	9703	ATTATACCATCGCCATCTGTGTACACCATCTTACCCCTGCATCAGCTACCGTGGCGGA	9762
OY	4563	TCTGTGTAGGCACTGCAATCATCAGACAGCTTGATCGCCCAAGCAAGAAAGATCGGCTGA	4622
Db	9763	TGATATATGGCGTAACTGTGACAGTGTATATGTGCTGTAAAGCGCCCTGAGTGCTGA	9822
OY	4623	CGCCATACGCGCTTTCACCGAAGCAAGATCCGATACAGACATTAAGCGGTTTGTGTGCA	4682
Db	9823	CGCCATAGCCCTGGCCCCCAACGCGCTAATCCCAATTCGCTGGGACCTCTTGCTGCG	9882
OY	4683	TTGGGCCAACCAACGCTGAACATTTGGAGAAACTTTGAACATCTGTGGTTTAACAACC	4742
Db	9883	TTAGTTCGGCCAAATCTGAACCGTTACACGAGAACCATGATTACTTGTTGTGCMAAGTC	9942
OY	4743	AAAGCTTCTCTGGGCGACAGTTGTGTCATTTCCCTGGGACGGCTGTGTAATTCGTTCCGCT	4802
Db	9943	AGCCCTTCTTGGGATCCAGTGTGCATATCTTTGGCGCTTTCATCTGTTAATCGCT	10002
OY	4803	GCTTTTACGTCTGATGCTTTTATTTATGTGTGACAGCGCTGCTGCGGGAAAGTATAGC	4862
Db	10003	GCTGCTCTGCTGCTGCTCTTTTATTTAGTGTGTGGCGCGCCCTACCTGGGAAAGTATAGC	10062
OY	4863	CTTTGGAACATGCGACCACTGTGCCAAATGTTCCGGGGATCCCGTATTAAGCGCTGGTGC	4922
Db	10063	CCTACGAACATGCGACCACTGTTCCAAATGTCCACAGATATACGTTATTAAGGCACTGTTG	10122
OY	4923	AAAGCGCAGTTAACCGCCGACCTTAACTGTGAGATATACGCTGCTCTATATGGAATTAAAC	4982
Db	10123	AAAGGGCAGGGATATGCCCGCTCAATTTGGAGATCATGTCAATGCTCTGGAGGTTTGC	10182
OY	4983	CTTCAACTTAACAAGAGTATCGTACCTGCAAAATTCACACAGTCAATCTTCAACCAAG	5042
Db	10183	CTTCCACCAACCAAGATATCATTAATCTGCAATTTACACATGTGGTCCCTCCCAAAA	10242
OY	5043	TTAAATGCTGCGGATCCCTCGAGTCAAGCAAGCATCTCTCAAAGCGGATTTACACATGCGCG	5102

10243	TCAAATGCTCGGGCTCCTTGGAAATGACGGCGGCGCTCATCGACACTATACCTGCAAGG	10302
QY	5103 TTTTGGCGGATGTGTACCTCTTATGTGGGGAGGCCCAAAATGCTTGTGTGACATGTAGA	5162
Db	10303 TCTTCGGAGGGGTCACCCCTTTATGTGGGGAGAGCGCAAAATGTTTGGCGACATGTAGA	10362
QY	5163 ACACACAACACTGAGTGAAGGCGCTACGTCGAGTTCGCTCCAGACTGCACATATAGATCAACGAG	5222
Db	10363 ACAGCCAGATGATGAGGCGCTACGTCGAAATTTGTCAGCAAAATTCGCGTCTGACACGCGC	10422
QY	5223 TCGCACTAAAGTTCAACACAGCTGCTCTGAAATCGGCGCTGCGTATAGTATACGGCAACA	5282
Db	10423 AGGCGATTAAAGTGACACACTGCGCGCATGAAGTAAGACTGCGTATTTGTATCGGGAAACA	10482
QY	5283 CCACCGCGCACCTGAGTACGTTTGTCAATGGCGTCAACCGCAGTTCTTACGGGACCTGA	5342
Db	10483 CTACCACTTTCTTAGATGTGTAGGTGAACGAGACTCAACACAGAAAGTCTTAAAGACTTGA	10542
QY	5343 AGGCTATAGACAGGGCCGATATACAGCGGCTTTTACCCCTTTCAGCCATAAGTGGATATCA	5402
Db	10543 AAGTCATAGCTGACACCAATTCAGCAATCGTTTACGCCAATTCGATCAATAAGTGGTATTC	10602
QY	5403 GAAAGGGGCTTGTTTTCAACTACAGACTTCCCTGATAGTATGAGCTATGAACACGAGAGCGT	5462
Db	10603 ATCGCGGCGCTGGTGTAACAATTATGACTTCCCGGAATATGGAGGATGAACACGAGAGCGT	10662
QY	5463 TCGCGCATTTCAAGCATCTCTCGCTGTATGCTACAGACATAGTAGCCCGCATGACATAC	5522
Db	10663 TTGGAGACATTTCAAGCTACCTCTTGTACTAGCAAGAGATCTCATCCGACACACAGCATTA	10722
QY	5523 GGCCTCGAAGCCCTTGTCAAGAACATCCACGCTCCCTTACACCCGAAGAGTATACGGGT	5582
Db	10723 GGGTACTCAAGCCTTCCGCCAAGAGCTGATCTCCGTTACAGCAGGCGCTCATACGAT	10782
QY	5583 ATGAAATGTGGAAGAAACAATCTACAGACGACCCCTGCAAGAAACAGCACCATTTGGATGTA	5642
Db	10783 TTGAGATGTGGAAGAAACAATCAGCGCGCCCACTGCAGAGAAACCGACCTTTGGGATGTA	10842
QY	5643 AAATTGAATGAGACCTCTGCGAGCGCTTAACTGTGCTTACGGGACATCCCTATCTCGA	5702
Db	10843 AGATTGCAGTAAATCCGCTCCGAGCGGTGAGCTGTTCAATACGGGAACATTTCCCATTTCTA	10902
QY	5703 TTGACATCCCTGATGACAGCTTTTGTGAGATCATCAACAATCAACCAATTTTAAAGTTA	5762
Db	10903 TTGACATCCCGAACCGCTGCTTTATACAGACATCAACATGACACACTGGTCTCAACAGTCA	10962
QY	5763 GCTGCACAGTACAGACATGTCATTTATTTCTGACACATTTGGTGGTTCCTTAACATTTACGT	5822
Db	10963 AATGTGAATGTCAGTGAATGACATTTATTCAGACACATTTCCGCGGGAATGGCCACCTTCAGT	11022
QY	5823 ACAAACTGACAGAGGAGGACATTTGTCAGTTCACTTCCACTGCAGACAGCTGTTTTGA	5882
Db	11023 ATGTATCCGACCGGGAAGTCAATGCCCGGTACATTCGATTCGACACACAGCAACTCTCC	11082
QY	5883 AGGAAGGACACACACATGTGACGTCCGTAGGACAGATTAACATACATTTTAAACAATCGA	5942
Db	11083 AAGGTGCAACATGATCACTGTCTGGAAGAAAGAACCGGTGACATGACACTTTAGACACGCGCA	11142
QY	5943 GCCCAAGGCAAAATTTTATATGTTTCGCTATGCGGCAAGAAAGTCCACTGCAATGCTGAAT	6002
Db	11143 GTCCACAGGCGAATCTTTATGCTATCGCTGTGTGGGAAGAAAGACAAATGCAATGCAAGAT	11202
QY	6003 GTAAACCAACGCGCGCAACCATATAATTTGGAAGAACCACTAAAGTGCACAAAGATTTCCAGG	6062
Db	11203 GTAAACCAACACGCTGACCAATATGCTGAGCAACCCGCACAAAATATGACAAAGATTTCAAG	11262
QY	6063 CGGCAGTTTCCAAAACATCTTGAACCTGCGCTGTGACTGTTTGGGGAGCATCATCCC	6122
Db	11263 CCGGCATCTCAAAAACATCAATGAGTTGGCTGTTTGGCCCTTTTCCGCGGCGCTCGTCCGC	11322
QY	6123 TCATTTGTTGAAGACTTATAGTGTGGTCTTGGAGCTCTAAGCTTATTAACAACAGTAAAT	6182
Db	11323 TATTAATTTATAGACTTATAGATTTTGTCTTGACAGCATGAATGCTGACTACAGACGAAAGT	11382

Oy	6183	GACTGACGGGGAACTACGA-CATGACGGTAAATCGATGACTTCCGAGGAAGCGTGT	6243
Db	11383	GACGGCTACGCCCCCATGATCCGACGACAAACTCGATGACTTCGAGGAAGTACTGT	11444
Oy	6242	GCATTAATGCCACGGCGCGCT	6261
Db	11443	GCATTAATGCATCAGGCTGT	11462
RESULT 10			
AAAT30787			
AAAT30787	standard; DNA; 16656 BP.		
XX	AAAT30787;		
XX			
XX	12-SEP-1996 (first entry)		
XX			
DE	Alphavirus-based eukaryotic layered vector pVGEV1S.		
XX			
KM	Alphavirus; Sindbis virus; vector; gene therapy; vaccine;		
XX	pVGEV1S; ss.		
OS	Chimeric Sindbis virus;		
OS	Chimeric synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	promoter	1..60	
FT		/*tag= a	
FT		/note= "wild-type Sindbis virus sequence	
FT		capable of initiating transcription"	
FT	CDS	60..5750	
FT		/*tag= b	
FT		/product= non-structural proteins NSP 1-3	
FT	CDS	5928..7579	
FT		/*tag= c	
-FT		/product= non-structural protein NSP 4	
FT	misc_feature	7579..7602	
FT		/*tag= d	
FT		/function= minimal junction region core	
FT	misc_feature	11647..11703	
FT		/*tag= e	
FT		/function= Sindbis polymerase recognition	
FT		sequence	
XX			
PN	W09617072-A2.		
XX			
PD	06-JUN-1996.		
XX			
XX	30-NOV-1995;	95WO-US15490.	
XX			
PR	15-MAR-1995;	95US-0405827.	
PR	30-NOV-1994;	94US-0348472.	
PR	18-JAN-1995;	95US-0376184.	
XX			
PA	(CHIR) CHIRON VIAGENE INC.		
XX			
PI	Belli BA, Chang SMW, Driver DA, Dubensky TW, Ibanez CE;		
PI	Jolly DJ, Polo JM;		
XX			
DR	WPI; 1996-277785/28.		
XX			
PT	New recombinant alpha-virus vectors - used to develop prods and		
PT	methods for use in gene therapy and in the prodn. of vaccines		
XX			
PS	Example 2; Page 186-194; 256pp; English.		
CC			
CC	A eukaryotic layered vector initiation system, pVGEV1S (AAAT30787),		
CC	is based on the Sindbis alphavirus. It is obtd. by inserting		
CC	the Sindbis clone pGSP6GENTrep into vector pCDNA3. The construct		
CC	can be used in the prodn. of Sindbis virus-based vectors used		
CC	to develop products and methods for use in gene therapy and in		
CC	vaccine prodn. pVGEV1S is capable of initiating alphavirus		

Dh 7723 GAGAGCAGGCGGCCCGATGCTGCCGCAAGGGCTGCTTCTCAAAATCCAGCACTGA 7782
Qy 2607 GAGGTGATAGTGAACCTTGACTTTCAAAACAGATCACTAAATCCCGCGCAGAGTCA- 2665
Db 7783 CCACAGCGGTGATAGTCCCTAGTCAATGAGCAGGCAATCAGACCTCAACCCCGAGCTCCAC 7842
Qy 2666 -----CCGCCAAGAGAGAGAGAGTGTCTCTAAGCCAAAACCTACTAGCCTTAAAAA 2720
Db 7843 GCCCGCCACCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7902
Qy 2721 AGAAGCAGCAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2777
Db 7903 AAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7962
Qy 2778 TGTGTATGAAGTGTGAGTTCGAGCAAGACATTTCCGATCATG---CTGAACCGCCAGTGA 2834
Db 7963 TGGCACTTAAGTTGAGAGCGCCAGAGATTTGTCAGCTCAAGAGAGAGAGAGAGAGATGCA 8022
Qy 2835 ATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2894
Db 8023 TCGGCGACGCACTGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8082
Qy 2895 TTGATATGAGCAATTAAGCGCGCGCTGAATTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2954
Db 8083 TCGACCAACCTGTCTATCAAAAGCTCAAAATTTACCAAGTGTGACAGATACAGACATGAG 8142
Qy 2955 ACGGCGACGTTCCCGCAGAGACATGAATTCAGACAGCTGCACTAACCCAGCGCAAAACAC 3014
Db 8143 TCGCAAGTTGCGCAGTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8202
Qy 3015 CGGCGCTTCAACACTGCGACCAAGCGCGAGTCCAGATGAGAGAGAGAGAGAGAGAGAG 3074
Db 8203 AAGGATTTCTAATTAATGCGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8262
Qy 3075 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3134
Db 8263 CTCGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8322
Qy 3135 TTGAGGCTATTTCTTGAAGAGTGCATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3194
Db 8323 TTGCGGATGATGCTCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8382
Qy 3195 CTTGGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3251
Db 8383 CTTGGAATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8442
Qy 3252 -----CACTAGTATCAGCGCTATGCTGCTTTCGAATGTCAGCTTCCATGCGACA 3305
Db 8443 CAGCACCACTGTCAAGGCAATGCTTGTCTCGAATGTCAGCTTCCATGCGACCGCC 8502
Qy 3306 CACCGGTGTCTATTTCACTGACCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
Db 8503 CGGCCACATGCTAATCCGCGAATCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8562
Qy 3366 ACAATCAAAATTAAGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
Db 8563 ACATATAGAGCTTATGATACCTGCTCAATGCTCAATTCGAGAGAGAGAGAGAGAGAG 8622
Qy 3423 CCAAGCAGAGATTAACGAGTACCTTCACTGACCAATGCTTCCATGCTGAGAGAGAGAG 3482
Db 8623 GCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8682
Qy 3483 ATTGCAAGACTCAACGCGCTGTTTCAAGCCCAATTAATAATGAGAACTGTGAGAGAGAG 3542
Db 8683 ACTGCCACATATGAAACCGGTTCAGCCCTGTATGATGAGACAGAGTGTGAGAGAGAG 8742
Qy 3543 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3602
Db 8743 CGGAGATTAACCATATGAGATACAGACTTCGCGCAATTTGATGATGATGATGATGATG 8802
Qy 3603 CTGCGGATGATCAACAAATTTCCGTTACATGCTTTTGCACAGAGAGAGAGAGAGAGAG 3662
Db 8803 CAGCAAGGCGAAACAGTACCGCTACATGTCGCTTAAGCAGATACACCGTTAAAGAG 8862

Qy 3663 ACAGTATGAGAAATAGTATGACAGCATCTGGAACCTGCGCTGCTTGGCCAAAG 3722
Db 8863 GCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8922
Qy 3723 GGTATCTTCTGTTAGTCAATGTCCTCAGGTGACAGTGAATCCGTGATGATGATGATGATG 3782
Db 8923 GATATCTTCTCTGCAAAATGCGCTCAGGGGACAGCTGAACGGTATGATGATGATGATGATG 8982
Qy 3783 GAGCATTTGAGATTTATGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3842
Db 8983 GCAACTGAGAGAGTCAATGATCACTGAGCCGCAAGATTAACCAAAATTCGTGAGAGAGAG 9042
Qy 3843 AGGAGTACTTGTTCACACCGCTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3902
Db 9043 AAAAATATGATTTACTCTCCCTTCAGGTAAAGAAATTCCTTGACAGTGTATGACAGCTG 9102
Qy 3903 TGAAGAGAGCTGTCCGAGTACATTAACATGACAGGCGAGCCCAACGCGATTAAGT 3962
Db 9103 TGAAA---ACAATGAGAGGCTATCATCTATGACAGGCGGAGCCGCAAGCTTATATCAT 9159
Qy 3963 CTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4022
Db 9160 CTTACTGGAAGATCATCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9219
Qy 4023 CTTACGAATGATTAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4082
Db 9220 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9279
Qy 4083 ACGGCTGCTATTAAGCAAAACAGTCAATTCCTTCAAGAGAGAGAGAGAGAGAGAGAGAG 4142
Db 9280 CTGTTGTCACCGGCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9339
Qy 4143 TCAACTGCGCGGATCTTATTAAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4202
Db 9340 TCAACTCAGCGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9399
Qy 4203 CATTCGCTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4262
Db 9400 CTTTCAAGTGTATCCCGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATG 9459
Qy 4263 AGTGTTCAAAGGATCAGCTTCCTCACTGATGATGATGATGATGATGATGATGATGATGAT 4322
Db 9460 ATGGCTTTAAACATCAGCTCCCAATTAATGATACAGACCACTTGAATGCTCACACCA 9519
Qy 4323 GAAATTTGGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4382
Db 9520 GAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9579
Qy 4383 ATTTTCTGTGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4442
Db 9580 ACTTCACTGTCAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9639
Qy 4443 TCTGGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4502
Db 9640 TCTATGCCAAAGAGTACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9699
Qy 4503 ACTATTTATCATCGGATCAGAGTATCACTGTCATGTCATGTCATGTCATGTCATGTCATG 4562
Db 9700 ATTAATCACTATCGCATCTGTGTATCACTTTAAGCGTCGATCAGCTGTCGAG 9759
Qy 4563 TCTGTGAGAGAGATGATCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4622
Db 9760 TGATGATTTGGGCTATCTGTCAGTGTATGTCGTATTAAGAGAGAGAGAGAGAGAGAG 9819
Qy 4623 CGCCATACGCGCTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4682
Db 9820 CGCCATACGCGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9879
Qy 4683 TTGCGCCAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4742
Db 9880 TTATGTCGCGCAATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9939

QY	4743	AACGGTTCTCTGGGCAAGTGTGGACATCTCTCGGAGGCGTTGTAATCTGTCCGCT	4802
Db	9940	AGCCGTTCTTCTGGGTCAGTTGTGCACACTTTTGGCCGGTTCATCGTTAAATCGCT	9999
QY	4803	GCTTTTATGCTGTCATCCCTTTTATTTGGTGTGAGCGCTCTGCTGGGAAAGTGTAGC	4862
Db	10000	ACTGCTCTGTGCTGCTCCCTTTTATGTGTGTGCGGCGCTACCTGGCGAAAGTGTAGC	10055
QY	4863	CTTTGAAACATGTGCACACTGTGTCAAATGTTCCGGGGATCCCGTATTAAGCGTGTGC	4922
Db	10060	CCTACGAAACATGTGCACCACTGTGTCCAATGTGTGCACAGATACGTTAAGGCACTGTGG	10119
QY	4923	AAGCGGAGGTTTACGGCGCACTTAACCTGGAGATCAAGCGCTGTCAATGCGGATTAACAC	4982
Db	10120	AAAGGCGAGGTTATGCCCCGCTCAATTTGGAGATCATGTCAATGCTCTCGGAGTTTTC	10179
QY	4983	CTTCAACTTAACAGAGTACGTGTGACTGCAAAATTCACACATCATTTCTTACACACAAG	5042
Db	10180	CTTCCACCAACCAAGAGTACATTAATCTGCAAAATTCACACATGTGTGTCCCTCCCAAAA	10239
QY	5043	TTAAATGCTGGGCTCCCTGAGTGCAGAGCATCTCAAAAGCGGATTAACATGCGCG	5102
Db	10240	TCAATCTGTGGCTCTCTTGGAAATGTACGCGGCGGCTCATGACACATTAACCTGCAGG	10299
QY	5103	TTTTTGGCGGTGTACCTTTTCAATGTGGGAGGCGGCACATGCTTCTGTACACATGAGA	5162
Db	10300	TCCTTGGAGGGGTGTACCCCTTTATGTGTGGAGAGGCGCATGTTTTGTGACAGTGAGA	10355
QY	5163	ACACACACTGATGTAGAGGCGTACGTGAGTTGCGTCCAGACTGCACATATGATCAGCAG	5222
Db	10360	ACAGCCAGATGTAGTAGGCGTACCTCAATTTGTACAGAGTTGCGGCTGTGACCAAGCGC	10419
QY	5223	TGCGACTTAAAGTTTCAACACAGCTGTGTGAAGTCGGCTGCGTATAGTATACGCGACA	5282
Db	10420	AGGCGATTTAAGTGTACACTGCGCGCATGAAGAATGAGACTGCGTATATGTATCGGAAAC	10479
QY	5283	CCACCGGCGACCTGTGATACGTTTGTCAATGGCGTCAAGCGATTTCTCACGGAGCTGA	5342
Db	10480	CTACCACTTTTCTTATGATGTGTACGTGAACGAGTACACACAGAGACGTCTTAAAGCTTGA	10539
QY	5343	AGGCTATGACAGGCGCATATCAAGCGCTTTTTCACCTTTGACCAATTAAGTGTGATCA	5402
Db	10540	AAAGTCAATGCTGTGACCAATTTTCAGCATGCTTTAGCGCATTTGATCAATTAAGGTGTATCC	10599
QY	5403	GAAAGGGGCTTGTTTACACTACGATCTCCCTGAGTATGAGCTATGAACACAGAGCGT	5462
Db	10600	ATCGCGGCTGTGTGTACATATGATCTTCCCGGAATATGAGGAGTAAACAGAGAGCT	10655
QY	5463	TGCGCGATTTTCAAGCATCTCTCGCTTATGTCTACAGCATATGATGCCCGCATGTACATAC	5522
Db	10660	TGCGAGCATTTCAAGCATCTCTCTTGACTGTGACAGAGTCTCATGTGCGACGACAGACATTA	10719
QY	5523	GGCGTGTGAAGCCCTTGTCAAGAACATTCACGCTCCCTACACCGAAGCATATCAGGGT	5582
Db	10720	GGCTATCTCAAGCCTTCCGCGAAGACGTGATGTCCGTATCAAGCAGGCGCATCAGGAT	10779
QY	5583	ATGAATGTGGAAGAACATCTCAGAGCAACCCCTGCAGAAACAGACCAATTTGATGTGA	5642
Db	10780	TTTGAGATGTGGAAAAAACAATCAGGCGCGCCATCTGCAGAGAAACCGACCTTTGGGTGTGA	10839
QY	5643	AAATTTGAATGTGAGACCTCTGTGAGCGCTTAACTGTGTGTTAGGGGACATCTCATCTCGA	5702
Db	10840	AGATTTGACATTAATCTCCCTCCGAGCGGTGTGACGTTCATCTCGGGAACATTTCCATTTCTA	10899
QY	5703	TTGACATCCCGATATGACGCTTTTGTGTGATCATCAGATTCACAAACAATTTTGAAGTTA	5762
Db	10900	TTTGACATCCCGAAGCGCTGCTTTATCAGACATATAGTGTACCACTGTGTCTCAACGTGA	10959
QY	5763	GCTGACAGATGACAGACTGATTTATTTCTGACAGCTTGTGTGTTCTTAACTATTAACGT	5822
Db	10960	AAATGTGAAGTCAGTGATGTGACTTTATTCAGACAGACTTCCGCGGGAGTGTGCCACCTGTCACT	11019
QY	5823	ACAAAGCTGACAGGAGGACATTTCTCAAGTTCACTTCCACTCCACGACAGCTGTTTGA	5882

Db	11020	ATGATCCGACCGGAAAGTCCAAATGCCCGTACATTGGCACTTGAGCACAAGCACTCTCC	11075
Qy	5883	AGGAAGGACACACATATGTACTCCGTAGCAGCATTAACATACATTTTATGACATTCGA	5942
Db	11080	AAGAGTGCACAGTACATGTCTCTGGAGAAAGACGGGTGACATACCTTTAGCACCGCA	11133
Qy	5943	GCCCAACAAGAAATTTTATAGTTTCCGTATGCGGCAGAAAGTCCACCTGCAATGCTGAAT	6002
Db	11140	GTCACACAGCGAACTTTATGCTATGCTGTGTGGAAAGAACATGCAATGCAGAAAT	11199
Qy	6003	GTAACACACCGCGCGACCATATATTGGAGAACACATATAAGTCAGCAAGAAATTCAGG	6062
Db	11200	GTAACACACACAGCTGACCATATGCTGAGCACTCCGACAAATAATGACAAAGATTTCAAG	11255
Qy	6063	CGGCAAGTTTCCAAAAACATCTTGGAACTGCTGCTTCACATGTTTGGGGAGCATCAATCC	6122
Db	11260	CCGCAATCTCAAAAACATCATGAGATGTGGTGTGTTTCCCTTTCCGGCGGCGCTCGTCCG	11319
Qy	6123	TCATTGTTGAGACCTTATAGTGTGGTCTGCAAGCTCTATGCTTATPAAACACAGTAGAT	6182
Db	11320	TATTAATTTATAGGACTTATGATTTTGTCTTGACAGCATGATGCTGACATCACAGAAAGAT	11379
Qy	6183	GACTGACCGGGGACACTGA-CATAGGGGTAAACTGATGTAATCTTCGAGGAAGCGTGGT	6241
Db	11380	GACCGCTACGCCCCCAATGATCCGACACAGCAAAACTGATGTACTTCGAGAGAACTGATGT	11439
Qy	6242	GCATTAATGCCACGCGCGCGCT 6261	
Db	11440	GCATTAATGCATCAGGCTGGT 11459	
RESULT 11			
AAAT0577	ID	AAAT0577 standard; cDNA; 11717 BP.	
AC	AAAT0577;		
DT	06-DEC-2000	(first entry)	
XX			
DE	Sindbis-like virus strain YN87448 complete genome.		
XX			
KW	Genome: Sindbis-like virus strain YN87448; primer: RT-PCR; vaccine;		
XX	epidemic; Sindbis encephalitis; evolution; epidemiology; ds.		
OS	Sindbis-like virus strain YN87448.		
XX			
PN	CN1252445-A.		
PD	10-MAY-2000.		
XX			
PF	27-OCT-1998; 98CN-0120694.		
XX			
PR	27-OCT-1998; 98CN-0120694.		
XX			
PA	(VIRO-) INST VIROLOGY CHINESE ACAD PREVENTIVE ME.		
XX			
PI	Liáng G, Zhou G, Li L;		
XX			
DR	WPI; 2000-443226/39.		
XX			
PT	Whole genome sequence of YN87448 virus strain and its cloning method -		
PS	Claim 1; Page 2-9; 24pp; Chinese.		
XX			

CC other epidemic Sindbis virus strains at the molecular level and is
CC useful for understanding the source, evolution and molecular
CC epidemiology of Sindbis viruses.

XX Sequence 11717 BP; 3300 A; 3109 C; 2924 G; 2384 T; 0 other;

Query Match 36.6%; Score 2370.6; DB 21; Length 11717;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 3737; Conservative 0; Mismatches 1999; Indels 77; Gaps 9;

QY 580 ACTGACGGATGAAACGGAGCGTATATTTCTCATGGGAAACAGGCCAAGCTACCTTC 639
DB 5758 ACTGACTACACCGGGGTGGGTACATATTTTCAGAGGACAGCGCCCTGGGCACTTGC 5817
QY 640 AACGAATCAGTACGTCAATGTAACTACAGAACTATTTGGATCGGGCGCTCATG 699
DB 5818 AAAAGAACTCGGTCGTGAGAACCACTTACAGAACCGACTTGAGGCAATTTCTGG 5877
QY 700 AGAAGTATACGCCCCCGCTCGATCTCGAAGAGAGAAATGTACAGAAAGAACTGC 759
DB 5878 AAGAAATCTACGCCCGGCTGCTGACACGTGAAAGAGAACACTCAATCAGTACC 5937
QY 760 AATTATCGCCTCTGAAGAAATAGAAAGCAGTATCAATCAAGAAAGTAAATATGA 819
DB 5938 AGATGATGCCACCGAAGCCAAACAAAGCAGTACAGTCTGAAAAGTAAACAGCA 5997
QY 820 AAGCAATTAACGCGACGACCTCTTCTGATTTGGGCAACATCTATCATCAAGTGA 879
DB 5998 AAGCAATTAACGCGACGACCTCTTCTGATTTGGGCAACATCTATCATCTGCCACAG-- 6054
QY 880 ATCTGTGAGTGTATACAGAGTCAATTAATCTGTACCAATCTACTCTGCAACGTAATTA 939
DB 6055 ATCCAGCAGATGTATTAAGATCACTACCGAAACCACTGATTTCCAGAGTATACAG 6114
QY 940 ACAGATTATCATCTGACAGGTCGCGTTAAACGTCACTTAATTCACAGAAATTT 999
DB 6115 CTAACTACTGACCCCAAGTTTCTGTAGCTGTTTGAACAATCTGTGATGAGAAAT 6174
QY 1000 ACCCTACAGTACAGTATTTGTATTAACAGTGAATTCAGTCTTATCATGATGTGG 1059
DB 6175 ACCGAGGTGACGATCTTATTCAGATCACGAGAGTACGATGCTTACTTGATATGGAG 6234
QY 1060 ACGGCGATCGTGTCTGATGATACAGCACTTTTTCGCGGTAACCTGAGAGCTACC 1119
DB 6235 ACGGGAAGTGTGCTTGTCTGATATCTGCAATTTTTCGCCCAAGTATGAAGTTACC 6294
QY 1120 CAAGAAGCATGCTATTTTTCAGCCAGATTAAGTACAGCTCCATCGCTTATACGA 1179
DB 6295 CGAAGACACGAGTATTAAGCCCAACATCCGAGTGCCTTCCATCAAGCATGACGA 6354
QY 1180 ATACATTACAAATGTATTTGGCTGAGTACTAAAGGAATTTGAAGTTACCAAAATGC 1239
DB 6355 ACAGTTGCAAAACGTGCTATTTCCCGACTAAAGAACTGACAGTCAACAAATGC 6414
QY 1240 GAGATTAACCTGTCTTATGATTCGCGGCAATTAATGTTGATTTTCAAGAAATACCAT 1299
DB 6415 GTGAAGTCCCAACACTGAGCTCAGCGACATTTCAAGTTGATGCTTTGAAAATATGAT 6474
QY 1300 GCAATGATGATCTGGGATACCTTTGCGGATTAACCTATTTGCGTAACTACAGAAACG 1359
DB 6475 GCAATGACGAGTATTTGGGAGAGTTTGGCCGAAAGCCAAATTAAGATCACTCTAGATTGG 6534
QY 1360 TTACGCAATATGTACAAAGTGAAGGGCCGAAAGCAGCATTTGTTGCCAAATCTC 1419
DB 6535 TTACCGCATATGTCGACGACCTGAAGGCGCTTAAGGCGCGCATGTTTCCAAAGACGC 6594
QY 1420 ATATATTAACACCTGTGAGAGATTAACATGATGATCAATTTGATGGATTTAAAGAG 1479
DB 6595 ATATATTTGTGCTCATTCGAAAGATGCTTATGATGATTTGCTATGACATGAAAGAG 6554
QY 1480 ATGTCAAGTATCTCCCGGACGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
DB 6555 ACGTGAAGTATACACTGTGACAGAAACACAGAAAGAAAGACGAAAGTATCAAGTATAC 6714

QY 1540 AGGTCGAGATCCCTTGTCTACCGCTTACCTTTGCGGATATCATCGGAATTAATGTCGTA 1599
DB 6715 AAGCCGAGAACCCCTGCGACCGCTTACTATGCGGATTCACCGGAGTTATGTGCGCA 6774
QY 1600 GACTGAATGGGTGCTTCTGCAATATTCATCTCTTTCGATGATGACGGGAATTT 1659
DB 6775 GGCTTACAGCGTTTGTCTACCAATTCATCAGACGCTCTTTGACATGTCCGAGAGACT 6834
QY 1660 TTGATGCGATTTATTTGCTGAACATTTCCACACGGCCCAACAGTATTTGAAACGCACTCG 1719
DB 6835 TTGATGCAATCAATGACGAACACTTCAAGAAAGTGAACCGTATCTGAAACGGAATTCG 6894
QY 1720 CGTGTGTTGAAAGGAGAGAGAGAGCTATTCCTCATTTTGGGCTGATGATCTTGAAG 1779
DB 6895 CCTGCTTGCACAAAGCAGAGAGAGAGCTATGCGTTAAACGGCCGTATATCTTGGAAAG 6954
QY 1780 ACTTAAAGTGTGACCAACCGCTCTTGAATTTGATTAAGAGCGGCGTTGCGCAATATCAT 1839
DB 6955 ACTTAAAGTGTGACCAACCGCTCTTGAATTTGATTAAGAGCGGCGTTGCGCAATATCAT 7014
QY 1840 CTGTGACCTTACCTAAGAGAACGAGTTTAAATTTGCTGATGATTAATCCGATGCT 1899
DB 7015 CCAACCATCTGCGCCACGCGGATACCGTTTCAAAATTCGGGGGATGATGAATTCGGAAATGT 7074
QY 1900 TCTTAAAGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGAGTACTACGTG 1959
DB 7075 TCTTAAAGCTGTTTGTCAACACACTAGTCAATATGCTGTTATTCGACAGAGATTTGAGAG 7134
QY 1960 AAGCGTTAACACGTCAGCGTGCAGCGCTCTATTCGCGCAAGATTAATAGTGCATGCTG 2019
DB 7135 AAGCGTTAAACGTCGCAATATGTCAGCATTTATTCGCGCAAGATTAATAGTGCATGCTG 7194
QY 2020 TCGTCTCCGACACCTTATGATGCGGAGAGATGCGCCACTTGTGGCTGAACATGGAAGTAA 2079
DB 7195 TAGATCTGACAAAGAAATGCTGAGAGGCTGACCACTGCTCAACCTGGAAGTAA 7254
QY 2080 TTATATGATGAGTATTTGATATCAAGACACCTTCTGTGGGGGATTTATCTGTGCTG 2139
DB 7255 TCAATGACGATGATGAGGAGAGAGACACCTTATCTTGTGGGATTTATCTGTGCAAG 7314
QY 2140 ACCAGATTAACAGGACACGCTGACAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
DB 7315 ATTGCGTTACTCTCACAGCGTGTGCGTGCAGACCTCTTAAAGGCTTTTAAAGCTTG 7274
QY 2200 GAAACCATTTGCGATGATGATACCAAGATCTGCGCCCGCGGCGCATGATGATG 2259
DB 7375 GTAAACCGCTCCAGCGGACGAGCAGACGAAAGACAGAAACGCGCTCTGATGATG 7434
QY 2260 AAGCAATGCGATGAACAGAAATTTGAAATTTACGAGCAGTATGTAAGGCTGATGATCCA 2319
DB 7435 AAGCAATGCGATGAACAGAAATTTGAAATTTACGAGCAGTATGTAAGGCTGATGATCCA 7494
QY 2320 GATACGAGATCATATGCGAGGCTGATCATCATGCTCTGTCCAGCTTAAGCCGAAAGCG 2379
DB 7495 GATACGAGATCATATGCGAGGCTGATCATCATGCTCTGTCCAGCTTAAGCCGAAAGCG 7554
QY 2380 TTAAAGATTTCAAGAGATTAAGAGGAGCCCATCACTCTACGCTGATGATGATGATG 2439
DB 7555 AAGAGCATTTCAAGCATCAAGAGGAGAAATTAAGGCACTCTTACGCTGATGATGATGATG 7614
QY 2440 TGACGTATGAGA-----CAGCAGCTTACCAACCGGAGAGATGTTTCCA 2482
DB 7615 CAGCATATGATATTTTATCTGATCAATATACACACACACACATGATTAAGAGATTTCT 7674
QY 2483 TACCTTACGCTGAATTTTCCACGATTTACCTTAACAATTCGATGCTTACGAGATCCA 2542
DB 7675 TTAAATATGCTGAGGCGCGCGCTTCCAGCCCGCATGTCATGATGAGGCGCGGAGAA 7734
QY 2543 AACCTTCTTACGCGCGCTGAGAGCGCTTTCGCGCCCGCTGCTGCTCAATTCGAAGAT 2602
DB 7735 GAGAGCAGGCGCGCGCTGATCTGCGCAGATGAGGCTGCTTCCCAATTCAGACACTGA 7794

QY 2603 CTTAGAGGTGATAGTCAACTTGACTTTCAAAC--GATCAGCTATCCGCCGAG 2660
Db 7795 CCACAGCGGTGAGTCCCTAGTCACTTGGACAGGCAATAGACTCAAAACCCAGCCAC 7854
QY 2661 GTCCACCGCCAAAGAAAGAAAGTGTCTTCTAACCAAACTTCTAGCTTAAAAA 2720
Db 7855 GCCCGCCCGCCGACAGAAAGAGCGGCCCAACCAACCGAAGCGAAAGAAACCA 7914
QY 2721 AGAAGCAGCAAGCCAAAGAGAGCAAAACG---AAGCTTAAACAGAGAAAGCAACGTA 2777
Db 7915 AAACACAGAGAAAGAAAGAAAGCACTTGCAGAAACCCAAACCCGAAAGAGACGCGTA 7974
QY 2778 TGTGTATGAAGTTGAGTCGGAACAAGATT--TCCGATCATGCTGTAACGCCCAAGTGA 2834
Db 7975 TGGCATTAAGTTGAGGCCGACAGACTGTTGAGTAAAAAATAGAGCGGAGATGTCA 8034
QY 2835 ATGATATGCTGCGCTTGTGAGAGAGGCTGATGAACCACTCCAGTTGAGAAAA 2894
Db 8035 TCGGCGACGCACTGCGCAATGAAAGAAAGTAAATAAACCCTCCACGTAAGAAAGCACTA 8094
QY 2895 TTGATATGAGCAATTAGCGGCCGCTGAATTGAAGGCTAGCATGTAGCACTTGAGT 2954
Db 8095 TTGACCACTCGTGTCTATTAAGCTCAATTCACCAAGTGTGACATATGCAATGAGT 8154
QY 2955 ACGCGACGTTTCCCGAAGACATGAATGACACGCTGAGTACACCAAGCAACAC 3014
Db 8155 TCGCAATGTGCGGTCAACATGAGAGAGGCGTTTCACTCAACAGTGAACACCTTG 8214
QY 3015 CGGCTTTTACAATGCGACCAAGCGGCGAGTCAAGTATGAGATGAGATTTACCGTAC 3074
Db 8215 AAGGTTTCAACAATGCGACCAAGCGGCGTCAAGTATGAGAGGCAATTTACATCC 8274
QY 3075 CGAGAGAGTGGGGGGGAAAGCGGACAGCGGAAAGCCATCTGTGACAAAGAGCAAG 3134
Db 8275 CCGCGAGATGAGAGGACAGAGAGACAGTGTGTCTCAATTAATGATTAATCAAGCGG 8334
QY 3135 TTGTGCTATTTGTTCTAGAGGTGCAAAATGAGGACGCGTACGCGCTTTCAGTGTCA 3194
Db 8335 TTGTGCGATAGTCTTGGAGGGGCTGATGAGGAAACAAGAACCCCTTTGCGTGTCA 8394
QY 3195 CTTGGAACCAAGAAAGGCTGACCAATTAGGAGTATCCCGAAGGTTCTGAACCGTGT--- 3251
Db 8395 CTTGGAATAGCAAAAGGAAAGCAATCAAGACAAACCCGGAAGGACAGAGAGTGTCTG 8454
QY 3252 -----CACTAGTTACAGCGCTATGCGCTTTCGAAATGTCACTTCCCATGCGCAAC 3305
Db 8455 CTGCACTAGTGTCAAGCGCATGTCTTGTCTGAAAGCGTGAAGCTTCCATGCAATGCC 8514
QY 3306 CACCGTGTGCTATTCATGACGCGGAGAACGAACTGACGCTGCTCGAAGAGAGCTCG 3365
Db 8515 CGCCCACTGTCTACACCGCGCAACCATCAAGAGCTTGTGACATCTTCTGAAGAGAGTGA 8574
QY 3366 ACAATCCAAATTTACAGACGCTGTGAGAAAGTGTGAATGTCAATC-----ACGCC 3419
Db 8575 ACAGAGAGGCTTACAGACCGCTGTCAAGCGCATTTGGGGTGGGATGTCTCGGCAAGA 8634
QY 3420 GCGCCAAACGAAGCTTACCGATGACTTCACTGACAGTCCCTTACTCGGGGTTTCTGCC 3479
Db 8635 GTAAAAAAGAGCTCTACCTGACCTTAACTTGAACAGCGCCGTAATCTTGGGCAATCTCT 8694
QY 3480 CGTATTTGAGACACTCAACGCGGTGTTTCAAGCCCAATTAATAATGAGAACGTTGGAG 3539
Db 8695 CGTACTGTACCATTAAGAACCGTCTTAACTTGAATGAGTCAAGAGGTTCTGGAGTGG 8754
QY 3540 AATGTATGATGATGATGATTAAGATTCAGGCTTGGCAATTCGCTTCAATCAATCAGGAG 3599
Db 8755 AAGCGAGAGCAAAACCATAGCATACAGACTTTCGCGCAAGTTTGGATACGCAAGAGG 8814
QY 3600 GCATCTGCGAGTGTCAAAATTCGTTACATGTCTTTGACAGCAAGCAATGACATCAAGG 3659
Db 8815 GAGCAGCAAGCTCAATATAGTACCGCTCATGATACCTGTGAGCAGATCATATCTGTCAAG 8874
QY 3660 AAGACATATGAGAAATAGCTATCAGACATGTGAGACCTGCGGTGTGTCGACACA 3719

Db 8875 AAGGACCATGATGATGATCAAAATCAGACACTGAGACCGGTGGAAGGCTTACTACA 8934
QY 3720 AAGGATTTCTCTGTAGTCAATGTCTCTCGAGGTGACAGTGTAAACGCTAGTATCAGA 3779
Db 8935 AAGGATTTCTCTCTCGCAAGTGTCTTCCAGGGGAGAGGTTACGCTTATGACATGGA 8994
QY 3780 GCGGACATCTGAGAAATTCATGACCGGTGAGAAATAATCATCGAGAGAGTGTGCGGTA 3839
Db 8995 GTAGCACTTCAAGCAAGTATGACATATGAGGCGCCGGAATTAACCAAAATTCGTGGAC 9054
QY 3840 GAGAGAGTACTTGTTCACCCGTCATGGAAGCTGTAAAGTCCACGTTTACGATC 3899
Db 9055 GGGAAAAATATGACTTACTCTCCGTTTCAAGAGATTCCTTGCAAGTGTGACACC 9114
QY 3900 ACTTGAAGAGACGCTTCCCGGTATCAATACATGACAGGCAAGGCCCAACGCGTATA 3959
Db 9115 GTCTGAAAAACAACCCCGGCTACATCACTATGACAGGCGCGGACCGCATATA 9174
QY 3960 AGTCTATCTGAGAGAGCGTCAAGCGAAGTATTAACCACTCTTGGCAAGAACG 4019
Db 9175 CATCTTATCTGAAGAAATCATGAGGAAAGTTTACGGAAGCCATTCGGAGAAACA 9234
QY 4020 TCACCTACGAATGTAAAGTGTGCGCATACAGACAGTATGTCAGCAAGCGAAGAA 4079
Db 9235 TTAGGTATGAGTGAAGTGGCGGATTAAGAGCCGGAACGTTTACGACCGGTACGAAA 9294
QY 4080 TGAAGCGCTGCACTTAAGCAAAACAGTCAATGCTTCAAGAGCGAACCAAAATGGG 4139
Db 9295 TCACGCGCTGACCGCCATCAAGAGAGTGTGCGCTTATTAAGAGCCAAACGAAGTGGG 9354
QY 4140 TCTTCAATCGCGGATTTATTAAGGACACAGACCACTCATGATGAAGTAAATTGACA 4199
Db 9355 TCTTCACTGCGCGACTCTATCAAGACGCGACACACGCGCTTAAGGAAATTTGCA 9414
QY 4200 TTCCATTCGCGTTGACACCGACAGTCTGCGCGGTTCCTTATGCTACACGCTTACGTA 4259
Db 9415 TGCCTTTCAAGCTATCCGAGTACTGTGATGCTGCTGCTTCCGACGCGGACGTAAG 9474
QY 4260 CGAAGTGTCAAAAGGATTCACCTTCACTGATGCAATGAGCAACATTTGCTGACAA 4319
Db 9475 TACACGCTTTTAAACATCAGCTTCCCAATTAAGACAGACCATCTGACATTTGCTACA 9534
QY 4320 CGAGAAATTTGGGGCTGAGACAGACAGACAGAGATTAACAGGCTTACATCA 4379
Db 9535 CGAGAGACTAGGGGCAAAACCGGAAACCAACCATGAATGATCATTCGAAAGAGGTTA 9594
QY 4380 GGAATTTTCTGTGGGGCGAAGAGGCTGAGATAGTATGGGGTAAACATGAAACGATCA 4439
Db 9595 GAACTTCAACGCTGACCGGATGCGGATGCTGAAATACATATGGGGCAATACGAACGATGA 9654
QY 4440 GAGTCTGGGCCAGAGTGGGACCAAGGCAACCAATGATGCGCGATGAGATCATCA 4499
Db 9655 GGGTCTATGAGCCAGAGTCTGACCAAGGAGCCCTCAAGATGGCCACAGAAATAGTAC 9714
QY 4500 TCCACTTATATTCGGGATTCAGTCTACATGTATGTGTGTGTGTGTGTGTGTGTGTGT 4559
Db 9715 AGCATTTACTATATTCGCTATCTGTGTACATCTTGAACCATTTGACCGTGTGTGTGT 9774
QY 4560 CTATCTGTGAGGACCTGATCATCATCAAGAGCTTGCATGCGCAAGCAAGAGAGATGCC 4619
Db 9775 CGATGATATGAGGATGATCTGTGACAGATTAATGTGCTGTAAAGCGCGGTGAGTGC 9834
QY 4620 TGACGCATATGCGCTTGACCGAAGCGAAGGATACCAAGCATTTAGCGGTTTGTGCT 4679
Db 9835 TGACGCCTATGAGCTTGGCCCAATGCGGATTCCTTCCGTTGCTGCACTTTTGTGCT 9894
QY 4680 GCATTCGCGCAACCAAGCGTGAACATTTGGAAGAACTTTGAACCATGTGTGTTTACA 4739
Db 9895 GTGTAGGTGTGGCTATATGTGAACATTCACCGAAGCATATGATTAATGTGTGTAACA 9954
QY 4740 ACCAACGTTTCTGTGGCAAGTTGTGATCTCTGTGGACGCGCTGTATTTCTGTCTCC 4799

Dh 9955 GCCACCGCTTCTTCTGGGTCAGCTGTGTATACCTCTGCGCGCTGTCTAATGTC 10014
Qy 4800 GCTGCTTTTATATGCTGATGCTTTTATTTATGTTGAGGCGTGTGCTGGGGAAGTAG 4859
Dh 10015 GCTGTCTATGCTGCTGCTGCTTTTATGTTGCTGGGCGCTTACTGCGGAAGTAG 10074
Qy 4860 ACCGCTTTCGAACATGCGACCACTGTGCGCAATGTTCCGGGAGTCCGTAATAGGCGTTGG 4919
Dh 10075 ACAGCTTTCGAACATGCGACCACTGTGCGCAATGTTCCGGGAGTCCGTAATAGGCGTTGG 10134
Qy 4920 TCGAACGCGAGGTTATGCGCGCACTTAACCTGAGATCAAGTGTCTCATGGAATTA 4979
Dh 10135 TTGAAAGGCGAGGTTATGCGCGCACTTAACCTGAGATCAAGTGTCTCATGGAATTA 10194
Qy 4980 CACCTTCACTAACAAGAGTACGTACCTGCAATTCACACAGTATTCCTTACACAC 5039
Dh 10195 TGCCTTCCACCAACCAAGAGTACGTACCTGCAATTCACACAGTATTCCTTACACAC 10254
Qy 5040 AAGTTAATGCTGCGGCTCCTCGAGTGCAGGACCTCTCAAAAGCGGATTACACATGCC 5099
Dh 10255 AAGTCAAGATGCTGCGGCTCCTCGAGTGCAGGACCTCTCAAAAGCGGATTACACATGCC 10314
Qy 5100 GCGTTTTTGGCGGTGTGATCCCTTTCATGTGGGAGCGGCAAACTCTTCTGTGACATG 5159
Dh 10315 AGGTCTTTTGGAGGCGGTGTGATCCCTTTCATGTGGGAGCGGCAAACTCTTCTGTGACATG 10374
Qy 5160 AGAACACACACATGAGTACGCGGTACGTCGAGTTCGCTCCGACATGCACTATAGATCAG 5219
Dh 10375 AGAACACACACATGAGTACGCGGTACGTCGAGTTCGCTCCGACATGCACTATAGATCAG 10434
Qy 5220 CAGTGCACATAAAAGTTTACACACAGCTGCTGAAAGTCCGCTGCTATATATATACGCA 5279
Dh 10435 CGCAGGCGATTAAGTGTGATCTGCGCGATGAAGTATGATCTGATATATATATACGCA 10494
Qy 5280 ACACACCGCGCACCTGGATACGTTTGTCAAATGCGCTCACCGCAGTTCTCACGGAAC 5339
Dh 10495 ACACATACATGTTCTCTAGATGTGTACGTGAACGAGTACACACCGAAGCTCTTAAGACC 10554
Qy 5340 TGAAGCTATGACGAGGCGGATATCAGCGCTTTTTCACCTTTCACCATTAAGTGTGCA 5399
Dh 10555 TGAAGCTATGACGAGGCGGATATCAGCGCTTTTTCACCATTAAGTGTGCA 10614
Qy 5400 TCAGAAAGGCGCTGTTTACAACTACAGCTTCCCTGAGTATGAGCTATGAACACAGAG 5459
Dh 10615 TCCATCGCGCTGTGTGATCACTATGACTTCCGGAATAGGAGCGATGAACACAGAG 10674
Qy 5460 CGTTCGCGATATTCAGACATCTCTGCTTGAATGCTACAGACATAGTACCGGCACTGACA 5519
Dh 10675 CGTTCGCGATATTCAGACATCTCTGCTTGAATGCTACAGACATAGTACCGGCACTGACA 10734
Qy 5520 TACGCTGCTGAAAGCTTCTGTCAAGAACATCCAGTCCCTTACACCCAAAGCAATATCAG 5579
Dh 10735 TTAAGCTATCAACCTTCTGCGCAAGAACGTGATGCTCCGTACACGAGGCGCATCTG 10794
Qy 5580 GGTATGAATGTGGAAGAACAACTCAGAGCAGACCTCTGCAAGAAACAGACATTTGGAT 5639
Dh 10795 GATTGAGATGTGGAAGAACAACTCAGAGCAGACCTCTGCAAGAAACAGACATTTGGAT 10854
Qy 5640 GTAAATTTGAAGTGAAGCTCTGCGAGCGTCTAACTGTGCTTACGGGCAATCCCTATCT 5699
Dh 10855 GCAAGATTTGAGTGAAGCTCTGCGAGCGTCTAACTGTGCTTACGGGCAATCCCTATCT 10914
Qy 5700 CGATTGACATCCGATGTCAGCTTTGTGATCATCTGAGATTCACCAACATTTTAAAG 5759
Dh 10915 CTATTGACATCCGATGTCAGCTTTGTGATCATCTGAGATTCACCAACATTTTAAAG 10974
Qy 5760 TTACTGACAGTACAGACTGCAATTTATCTGAGACTTGTGTCTTAAACATTTAC 5819
Dh 10975 TCAATGTGATGTGAGTGTGCACTTATTCAGGGAATTTGGGCGGATGTGCTACCTGC 11034
Qy 5820 AGTACAAAGCTACAGGAGGACATGTTCAGATTCACTCCACACGACAGCTGTTT 5879
Dh 11035 AGTATGTATCGACCGGAGGACATGTTCAGATTCACTCCACACGACAGCTGTTT 11094

Qy 5880 TGAAGAGGACACACATGATGATGCTCCGTAGGACAGATACATATTTAGACAT 5939
Dh 11095 TCCAGAGTGTGACAGTTCATGTCTGTGAGAAAGAGGCGTGTGATCTTACGACCG 11154
Qy 5940 CGAGCCCAAGCAAAATTTATGATTTGCTATGCGGCAAGAAAGTCACTGCAATGCTG 5999
Dh 11155 CGAGCCCAAGCAAAATTTATGATTTGCTATGCGGCAAGAAAGTCACTGCAATGCTG 11214
Qy 6000 AATGTAAACACCGGCGGACACATTAATGGAACACACATTAAGTGTGACCAAGATTC 6059
Dh 11215 AATGTAAACACCGGCGGACACATTAATGGAACACACATTAAGTGTGACCAAGATTC 11274
Qy 6060 AGGCGGAGTTTCCAAAACATTTTGAACGTGCTGCTTGGACCTGTTGGGGAAGCATCAT 6119
Dh 11275 AAGCGGAGTTTCCAAAACATTTTGAACGTGCTGCTTGGACCTGTTGGGGAAGCATCAT 11334
Qy 6120 CCTCATGTTGTGAGGACTTATAGTGTGTGCTGACGCTTATGCTTATTAACACAGCTA 6179
Dh 11335 CGCTATTAATTAATGATGATTTTCTTGCAGACATGATGCTGATGACACAGAA 11394
Qy 6180 GATGACTGAGCGGGAACATGCA-CATAGCGTAAACTGATGATCTTCCGAGGAAGCT 6238
Dh 11395 GATGACTGAGCGGGAACATGCA-CATAGCGTAAACTGATGATCTTCCGAGGAAGCT 11454
Qy 6239 GGTGATATATGCAAGCGCGCT-----TGAC 6265
Dh 11455 TGTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11514
Qy 6266 ACTAAACCTGATGATATTTCCGAGGAAGCAAGTCAATATGCTGTGACGT 6318
Dh 11515 ACCAAACTGAGTATTTCCGAGGAAGCAAGTCAATATGCTGTGACGT 11567

RESULT 12
AAV60125
ID AAV60125 standard; DNA; 16656 BP.
XX
AC AAV60125;
XX
DT 25-MAR-2003 (updated)
DT 04-DEC-1998 (first entry)
DE
XX Representative eukaryotic layered vector initiation system.
XX
KM Eukaryotic layered vector initiation system; stimulate; immune response;
KM sb.
OS Synthetic.
OS Sindbis virus.
XX
PN US814482-A.
XX
PD 29-SEP-1998.
XX
PF 30-OCT-1996; 96US-0739158.
XX
PR 14-SEP-1994; 94NO-US10469.
PR 15-MAR-1995; 95US-0404796.
PR 15-SEP-1993; 93US-0122791.
PR 18-FEB-1994; 94US-0198450.
PR 30-NOV-1994; 94US-0348472.
PR 18-JAN-1995; 95US-0376184.
XX
PA (DRIV/) DRIVER D A.
PA (DUBE/) DUBENSKY T W.
PA (JOLLY) JOLLY D J.
PA (POLO/) POLO J M.
XX
PI Driver DA, Dubensky TW, Jolly DJ, Polo JM;
XX
DR WPI; 1998-541753/46.
XX

PT Eukaryotic layered vector initiation system - containing eukaryotic
PT promoter and heterologous antigen coding sequence, useful for
PT stimulating immune response

PS Example 1; Fig 3A-H; 144bp; English.

CC The present sequence is a representative eukaryotic layered vector
CC initiation system, based on Sindbis. The eukaryotic layered vector
CC initiation system comprises a eukaryotic promoter 5' of viral cDNA
CC which initiates, in a susceptible target cell, 5' to 3' synthesis of
CC RNA from the viral cDNA. The RNA comprises a vector construct which
CC autonomously amplifies in the cell and expresses a heterologous nucleic
CC acid sequence which encodes an antigen or modified form that stimulates
CC an immune response within an animal. The system is useful for stimulating
CC an immune response to an antigen in vivo, especially where the immune
CC cells, preferably by infection in vivo, introducing the vector into target
CC response is a cell mediated, HLA class I-restricted or an HLA class
CC II-restricted immune response.
CC (Updated on 25-MAR-2003 to correct PR field.)

CC Sequence 16656 BP; 4467 A; 4349 C; 4183 G; 3657 T; 0 other;

Query Match 36.5%; Score 2369.2; DB 19; Length 16656;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3693; Conservative 0; Mismatches 1983; Indels 44; Gaps 9;

QY	580	ACTGACGGTATGACGGGAGGTATTTCTCATCGGAACAGGCCAAGTCACTTC	639
DB	5746	ACTGACTAACCGGGGTAGGTGGTACATATTTTCACGACACAGCCCTGGGCACTTGC	5805
QY	640	AACAGAAATCACTAGCTCAATGTAACTACAGAACTTATTTGGATCGGCGCTCATG	699
DB	5806	AAAAGAGTCCCTTGTGCAAGACCACTTACAGAACCACTTGGAGCCCAATGTCCTGG	5865
QY	700	AGAAGTATTAACGCCCCCGCTCGATCTTCGAAAGAGAGAAATTTTACAGAAAGATTCG	759
DB	5866	AAAAGATTCATGCCCCGGTCTCGACACGTCCAAAGAGAGAACACTCAAACTCAGATAC	5925
QY	760	AATTATGGGCTCTGAAGGAATATAGAGAGATTCATCAAGAAAGTAAAGTAAATATGCA	819
DB	5926	AGATATGCCCCACCGAAGCAAAAGAGTACAGTCTGTAAAGTAAAGTAAATATGCA	5985
QY	820	AAGCAATTAACAGCGAGCACTATTTCTGATTTGGGCACATATCATCAGAAATGA	879
DB	5986	AAGCATTAACCACTAGGAGCACTACTGTACGACTACGACTGATTAACCTTGCACAG---	6042
QY	880	ATCTCTCGAGTGTATACAGAGTCAATTTATCTGTACCAATCTACTGTCACAGGTAATTA	939
DB	6043	ATCACCGAAGATGTATTAAGATCACCTATCCGAAACATTTGATCTCAGTAGCGTACCGG	6102
QY	940	ACAGTTTACATCTGCAAGAGTCCGGGTTAAACGTGCACTTATGATTCAGAGCAATT	999
DB	6103	CGAACTACTCCGATCCACAGTTCGCTGTAGCTGTGTAAACAATCTGCTGAGTGAACCT	6162
QY	1000	ACCCTACAGTACCGCACTTATTTGATTAACAGATGAATAGATCGTATCTGACAGTGG	1059
DB	6163	ATCCGACAGTACATCTTATTCAGATTTACTGACGAGTACGATCTTATCTGATATGATG	6222
QY	1060	ACGGGCACTCGTGTCTGATAGTACAGCACTTTTGTCCGGCTAAACTGAGAACCTAC	1119
DB	6223	ACGGGACAGTCCGCTGCTGATCTGCAACTTCTGCCCCGCTAAAGTTAAGTTAC	6282
QY	1120	CAAAAGACATAGCTATTTTGCAGCAGAGATAGATCAGCGCTCCCTATGCTTATACGA	1179
DB	6283	CGAAAAAATATGATAGAGCCCGGAATATCCGAGTCCGGTTCCATCAGGATGACGA	6342
QY	1180	ATACATTAACAAATATTTGCTGCAAGTAAAGAAATGGCAAGCTTACCAAAATGC	1239
DB	6343	ACACCTTAACAAATATGCTCTATTTGCGCAACTAAAGAAATGGCAAGCTACGAGATGC	6402
QY	1240	GAGAAATACCTGTCTTAAGTTGCGCGGCAATTTAATGTTGATGTTTCAAGAAATACGAT	1299
DB	6403	GTAATCTGCAACACTGAGACTCAGCGACATTCATATGTGATGCTTTGCAAAATATGCT	6462

QY	1300	GCAATGATGAGTACTGGGATACCTTTCGCGATTAACCTTATTCGGGCTAACTACAGAAACG	1359
DB	6463	GTAATGACAGATTTGGAGAGAGTTCGCTCGGAACCAATTTAGATTCACATGAGTTTG	6522
QY	1360	TTACGCAATATGTGCAAAAGCTGAAAGGCGGAAAGCAGACAGATTTGTTGCAATATCTC	1419
DB	6523	TCACGCGATATGATCTAGTACGATGAAGGCCCTTAAGCCGCCCACTATTTGCAAAAGCT	6582
QY	1420	ATAATCTAAACCGTTGCAAGAGATACCAATGATCAATTTGCTCATTCGATCTTAAAGACG	1479
DB	6583	ATAATTTGCTCCATTTGCAAGAAAGCTTATGATGATTTGCTCATTTGCAATGAAAGAG	6642
QY	1480	ATGCAAAATTAATCTCCGCGACGAAACATACAGAGAGCGGCTAAGGTCAGGTTATTC	1539
DB	6643	ACGTAAAGTTTACACCAAGCAGCAAGAAACACAGAAAGAAAGACGAAAGTACAAATGATAC	6702
QY	1540	AGGCTGCAAGATCCCTTCTACCGCTTACCTTTGCGGATCCATCCGGAATTAATGTCCTA	1599
DB	6703	AAGCGGAGAAACCCCTGGGAGTCTGCTTATATGCGGATTCACCGGGAATTAATGTCCTA	6762
QY	1600	GACTGAATGCGGTCTTTCGCCAAATATCATCTCTTTCGACATGTCAGCGGAATG	1659
DB	6763	GCTTACCGCGCTCTTCTTCCAAACATTTCAACCGCTTTTGAATGTCGCGGAGATG	6822
QY	1660	TTGATGCAATTAATGCTGAACATTTCCACACCGGAGCCCGATTTGGAACGAGCATCG	1719
DB	6823	TTGATGCAATTAATGCTGAACATTTCCACACCGGAGCCCGATTTGGAACGAGCATCG	6882
QY	1720	CGTCTGTTGATTAAGCAAGACGACGCTATCCGCTTTCGCTTGAATGATCTCTGAGG	1779
DB	6883	CATCATTTGACAAAGCAAGACGACGCTATTCGCTTGAATGATCTCTGAGG	6942
QY	1780	ACTTAGTGTGACCAACCGCTCTTGAATTTGATAGAGCGCGCTTGCAGAAATATCAT	1839
DB	6943	ACTGCGGTGTGATCAACACTACTGCACTTGAATGATGATGCGCTTTCGAGAAATATCAT	7002
QY	1840	CTGTGCACTTCTCTACAGAAACGAGGTTAAATTTGGTCCATGATGAAATCCGCTATGT	1899
DB	7003	CCACCATCTACTACGAGTACTGCTTTTAAATTTGGGCGATGATGAAATCCGAAATGT	7062
QY	1900	TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTGACAGATCACTACGTG	1959
DB	7063	TCTTCACTTTTGTCAACACAGTTTGAATGTCATATGCGCCAGAGATCACTAGAAAG	7122
QY	1960	AAAGGTTAACCACTGACGCTGCGCGCTCTATTCGCGAGCAATTAAGTGCATGCTG	2019
DB	7123	AGCGGCTTAAACGTCAGATGTGACAGCTTCAATTTGCGAGCAACATCAATTAATGAG	7182
QY	2020	TCTGTCTCCGACCTTGAATGCGGAGAGATGCGCCACTTGGCTGAACATGGAAGTAA	2079
DB	7183	TAGTATCTGAACAAAGAAATGCTGAGAGTCCGCTCACTGAACATGAGGTTAAGA	7242
QY	2080	TTATGATGACGTTATTTGATCAAAAGCACTTCTGTTGGGGGATTTATCTCGTGG	2139
DB	7243	TCAITGACGCACTATCGGTAGAGACCACTTACTTCTGCGGCGAATTTATCTTTCGACG	7302
QY	2140	ACAGATTAACAGGACAGCTTCGACAGATGCGACACCTCTTAAAGGCTTTTAAAGCTTG	2199
DB	7303	ATTCGTTACTTCCACACAGCTGCGCGCTGCGGAGATCCCTGAAGGCTGTTAAGTTGG	7362
QY	2200	GAAGAACTTTGCGATCGATATACCAAGATGCGACCCGCGCGGCACTGCAATGATG	2259
DB	7363	GTAACCTGCTCCAGCCGACAGACAGAGCAAGAAAGAGAGAGCGCTCTGCTAATG	7422
QY	2260	AAGCAATGCGATGGAACAGAAATTTGAATTAAGSAGATTAAGTGAAGGCGGTAGAAATCA	2319
DB	7423	AAACAAAGCGGTGTTTGAATGATGATTAACAGGACTTTTGAAGTGCAGTGCAGAGATCC	7482
QY	2320	GATACGAGATCAATCTGGCAGCGCTGATCATCGTCTCTGTCCAGCTTACCCGAAACCG	2379
DB	7483	GGTATGAGTGAACAATTTACACCTGCTCACTAGGCAATGAGAACTTTTGCACGAGCA	7542

Oy	2380	TTTAGAATTCAAGAGCAATAAGAGGAGACCACCATACCCCTTAGCGCTAGACTTAATATGG	2439
Db	7543	AAAGAGCATTCACAGCCATCAAGAGGGGAATAAAGACATCTTAGCGGTGCTCTAAATAGT	7602
Oy	2440	TGACGTAGTAACAAGCACTAC-----CCAGCGGAGAAATGTTTCCATACC	2486
Oy	2487	CTCAGCTGAACCTTTCACACAGTTTACCTTACAAATCCGATGGCTTACCAGATCCAAAC	2546
Db	7663	TTAACATCTCGGCGCGCCCTTCCGGCCCCCATGCAATGGAGGCGGAGGAGA	7722
Oy	2547	CTCTAGCGCGCCGTGAGGCCGCTTGGCCCCCGCTGGCTGCTTAAATCGAATCTTA	2606
Db	7723	GGAAGCAAGCGGCCCCGATGCTCCGCCCAACCGGCTGGCTTCTCAATATCAGCACTGA	7782
Oy	2607	GGAGGTGATATAGTCACTTTGACTTTCAACAACGATTCACATATCCGCGCGCAGGTCCA	2665
Db	7783	CCACAGCCGTCAATGCCCTTACTGATTTGAGCAGGCACTAGACCTTAACCCCACTCCAC	7842
Oy	2666	-----CCGCCAAAGAAAGAAAGATGCTCTTAAGCCAAAACCTAAGCTTAAGAA	2720
Db	7843	GCCGCCACCGGGCCAGAAAGAGCGGGCCCAAGCAACACCGAAGCCGAAGAAACCA	7902
Oy	2721	AGAAAGCAACCAAGAGAGCAAAAC--GCAAGCTTAAACCAAGGAAAGCAACGTA	2777
Db	7903	AAACCGCAGAAAGAAAGAAAGAACCACTGCMAAACCCMAACCCGGAAGAGACAGCGCA	7962
Oy	2778	TGTGTATGAATGGAGTCCGGAACAAGACATTTCCGATCATG---CTGAAGCGGCCAAGTGA	2834
Db	7963	TGGCACTTAAGTTGAGGCGCAAGATTTGTTGACGTACAGAACGAGACGGAATGTCA	8022
Oy	2835	ATGATATATGCTTCGTTTCGAGAGAGAGCTGATGAAAACACTCCACGTTGAAGAAAA	2894
Db	8023	TCGGGCAACGACTGGCCATGGAAGAAAGTAATGAAACCTCTGCACGTGAAGAAACCA	8082
Oy	2895	TTGATATAGAACATTTACGGCCCGTGAAATTGGAAGAGCTTAGCATGTACGACTTGGAGT	2954
Db	8083	TGCACACCTCTGTGCTATCAAAAGCTCAAAATTTTCCAAATCTCTCAGCAATACATGAGT	8142
Oy	2955	ACGGGACGTTCCCCAGAAACATGAATACAGACAGCTCAGTACACAGCGACAAACAC	3014
Db	8143	TCGCACAGTTGCACTGCAATGAGAAAGTGAAGCACTTACCTTACCAAGTAGACACCCCG	8202
Oy	3015	CGGGCTTCAACACTGGACCAACCGCGCAGTCCAGTATGAATAAGGAGATTTACCCGTAC	3074
Db	8203	AAGATTTCTAATACGCGACCAACGAGGGGTGAGTATAGTGAAGTGAATTTACATCC	8262
Oy	3075	CGAGAGAGTGGCGGGAAGCGCACAGCGGAAGACGATCTCTGGAACAACAGAGCAGAG	3134
Db	8263	CTCGGGGAGTGAAGGACAGAGAGACACCGGTCTGCATCATGATTAATCTCCGTCCGG	8322
Oy	3135	TTGTGGCATTTGTTCTAGAGGTGCAATGAAGGGGACGCGTACCGCGCTTCAATGGTCA	3194
Db	8323	TTGTGCGCATATGCTCTCGGTGGCGCTGATGAAGAAACGAACCTGCCCTTTGGGTCTCA	8382
Oy	3195	CTTGGAAACGAAGAGGGGTGACATTAAGGATACCCCGAAGGTTCTTGAACCGTGT---	3251
Db	8383	CCTGGAAATAGTAAGGAAGACATTTAAGACGACCCCGGAAGGGGACAAAGATGTGTCCG	8442
Oy	3252	-----CACTAGTTACAGCGGCTATGCGTCTTTGCAATGTCAAGTTCCCATGCGCAAAAC	3305
Db	8443	CAGCAACCACTGGTCAACGCAATATGTTGTCTCGAAATATGTAAGCTTCCCATGCAACGCG	8502
Oy	3306	CACCCGTGTGCTATTCACCTGACGCCAAGACCACTCGACGTGTGCAAGAAAGCTCG	3355
Db	8503	CGCCCAATGCTAATACCGCGAAACCTTCAAGCGCTCGACATCTTTGAAAGAAACGTA	8562
Oy	3366	ACAATCCAAATTAACGACACGCTGTGAGAAAGCTTTGAATGTGTCA---TCAAGCGCGC	3422
Db	8563	ACCATAGAGCCTAGATACCTGTCTCAATGCGATTTTGGGTGGGATGCTGTGGCAGAA	8622
Oy	3423	CCAAACGAAGCATTAACGATGACTTCAACTGACCAATGCTCCATCTGGGGTTTGTCCCGT	3482

[illegible]

Dh 9700 ATTACTACCATGCGCATCTGTGTACACCATCTTACCGGTGCGCATCGACTACCGTGGCA 9759
Qy 4563 TCTGTAGGACACTGCATCATCAGCAGGCTTGATCGCCAAAGCAGAAAGAGACTGCTGA 4622
Db 9760 TGAATATTGGCTAATCTGTGACAGTGTATGTGCTGTAAAGCGCGGTGATGCTCGA 9819
Qy 4623 GCGCATACGCGCTTGACCGAAACGCAAGGTAACCCACAGCATTTACGGTTTGTGCTCA 4682
Db 9820 GCGCATACGCGCTTGACCGAAACGCGTAATCCCAACTTCGTTGCGACTTTGTCTGCG 9879
Qy 4683 TTGCGCAACCAACGCTGAACATTGGAGAACTTTGAACCATCTGTGTTAAACAAC 4742
Db 9880 TTAGTGTGGCCCAATCTGAATGACCGAGACATGATTAATCTGTGTGAAACAGTC 9939
Qy 4743 AACCGTTTCTGGGACAGTTGTGATCTCTGCGACGCGTGTATTCGTGCTCGCT 4802
Db 9940 AGCGTTTCTTGGGATCGATTTGTGATACCTTTTGCGCGGTTCATCGTTCAATGCGCT 9999
Qy 4803 GCTTTTCAATGCTGACATGCTTTTATTTGTTGACAGCGCTTGCTGGGAAAGTACG 4862
Db 10000 ACTGTCCTGCTGCTGCTGCTTTTATTTAGTGTGGCGGCGCTTACTGGGAAAGTACG 10059
Qy 4863 CTTTGAACATGCGACCACTGTGCGCAATGTTCCGGGATCCGTAATAAGCGTGTGCG 4922
Db 10060 CTTAGGAACATGCGACCACTGTTCCAAAATGTCCACAGTACCGTAATAAGGCACTTGTG 10119
Qy 4923 AACCGGAGGTATGCGGCACTTAACCTGAGATACGCTGCTCATCGAATTAACAC 4982
Db 10120 AAAGGCGAGGATCCCGCTCAATTTGAGATCACTGTCAATGCTCTGGAGGTTTTCG 10179
Qy 4983 CTTCAACTAACAGAGATGACGTGACCTGCAAAATTCACACAGTCAATTCCTTACACCAAG 5042
Db 10180 CTTCCACCAACAAAGATACATTAATCTGCAATTCACACATGCTGCTCCCTCCAAAAA 10239
Qy 5043 TTTAAATGCTGGGGTCCCTCGAGTCAGAGCATCTCTAAAGCGGATTAACACATCCGCG 5102
Db 10240 TCAATGCTGGGGTCTCTTGGATGTCAAGCCGCGCTCATATGCACTATACCTGCAAGG 10299
Qy 5103 TTTTGGGGGTGTATACCTTTTATGTTGGGAGGCGCAATGCTTCTGTGACATGAGA 5162
Db 10300 TCTTGGAGGGGTCTAACCCCTTATGTGGGAGGCGCAATGTTTTCGCACTGTAGA 10359
Qy 5163 ACACACAATGAGTAGGCGGTACGTGAGTTCGCTCCAGATGCACTATAGTACAGCAG 5222
Db 10360 ACAAGCAATGATGATGAGCGGTACGTGATTCAGCAGATGCGCGCTGACACGCGC 10419
Qy 5223 TCGCACTAAAGTTCAACAAGTCTGTGAAGTGGCGCTGGTATAGTATACGGCAACA 5282
Db 10420 AGGCAATTAAGGTGACACTGCGCGAGTGAAGTGAAGTGTATGTTACGGGAACA 10479
Qy 5283 CCAACGCGCACCTGATACGTTTGTTCATGAGGTGACAGCGAGTTCCTACGGGACTGA 5342
Db 10480 CTACCAATTTCTTAATGTATGTAAGTGAACGAGTACACAGGAACGTTCTAAAGCTTGA 10539
Qy 5343 AGGTATAGCAGAGCGGATATCAGCGGCTTTTTCACCTTTGACCATTAAGGTGCTATCA 5402
Db 10540 AAGTATAGCTGAGCAATTTTCAGATGTTTACCCATTTGATATTAAGTTCGTTATTC 10559
Qy 5403 GAAAGGGCTTGTTTTAACTACGATTCCTGATATGAGCTATGAAACCGAGAGCGT 5462
Db 10600 ATCGGGCTGTGTATCAACTATGACTTCCCGAATATGAGAGCATGAACAGAGAGCGT 10659
Qy 5463 TCGGGAATTTCAAGATCTCTGCTTGAATGCTACAGACATAGTACCGGCACTGACATAC 5522
Db 10660 TCGGGAATTTCAAGATCTCTCTTGAATGCTACAGACATAGTACCGGCACTGACATTA 10719
Qy 5523 GCGTCTGAGAGCTTCTGTCAAGACATTCAGCTCCCTTACACCAAGCAAGATACAGGT 5582
Db 10720 GCGTATCTCAAGCTTCCGCAAGAACGTGCAATGCTCCGTACACGAGCGCATACAGAT 10779
Qy 5583 ATGAATGTGAGAGAACCACTCAGAGCAAGCTTGTCAAGAAACAGCACATTTGGATGTA 5642
Db 10780 TTGAGATGTGAGAAACCACTCAGAGCGCGCTGCAAGAAACCGCACCTTTGGGTGTA 10839

Qy 5643 AAATTGAAGTGAAGCTCTGCGAGGCTTAATGTTGCTTACGGGACATCCCTATCTGCA 5702
Db 10840 AGATTGCAATTAATCCCTCCGAGCGGTGACCTGTTCTATAGGGAACATTTCCATTTCTA 10899
Qy 5703 TTGACATCCCTGATGACACTTTTGTGAGATCATCAGATCAACAATTTTGAAGTTA 5762
Db 10900 TTGACATCCGAAAGCTCTCTTATACAGACATGAGATGACACACTGCTTCAACAGTCA 10959
Qy 5763 GCTGACAGTGAAGACTGCACTTATTTGTGACAGATTTGTGTGTTCTTAACATTAAGT 5822
Db 10960 AATGTGAAGTGAAGTGAAGCACTTATTCAGACAGCTTGGCGGAGTGGCCACCTGCACT 11019
Qy 5823 ACAAGCTGACAGGAGGAGCAATTTGCTGCTTCACTCCACCTCCACGACAGCTGTTTGA 5882
Db 11020 ATGTATCCGACCGGAGAGGTAAATGCCGTACATTCGCAATTCGACACAGCACTTCC 11079
Qy 5883 AGAAGCGACCAACATGTGACTGCGGTAGGACAGATTAACATTAATTTAGACATGCA 5942
Db 11080 AAGAGTGAACAGTACATGCTCTGGAGAAAGAGCGGTGACAGTACATTTAGCACGGGA 11139
Qy 5943 GCCCAAGCAAAATTTTATGTTGCTATGCGGACAGAAATCCACTGCAATGCTGAAT 6002
Db 11140 GTCCACAGCGCACTTATGCTATGCTGTGTGGAGAGAAACAACATGCAATGCAAGAT 11199
Qy 6003 GTAAACCAACGCGCGCACATTAATTTGAGAAACAACAATAAGTGCACCAAGATTCACAG 6062
Db 11200 GTAAACCAACGCTGACCATATTCGAGACACCCGCAACAATAAGCAAGATTTCAAG 11259
Qy 6063 CGGAGTTTCCAAACATCTTGAACCTGCTGCTTGAACCTTTGGGGAGACATCATCC 6122
Db 11260 CCGCATCTCAAAAACATCAAGAGTGTGCTGTTGCCCTTTTGGCGCGCTGCTGCG 11319
Qy 6123 TCAATTTTGAAGCTTATGATGTTGTCTGACGCTATGCTTATTAACACAGTAGAT 6182
Db 11320 TATTAATTAATGAGACTTATGATTTTGTCTGACATGATCTGACACAGAGAT 11379
Qy 6183 GACTGAGCGCGGACACTGA-CATAGCGTAAACTGATGTAATCTCGAGGAAGGTGCT 6241
Db 11380 GACCGCTACGCCCAATGATCGACAGCAAAACTCGATGTAATTCGAGGAACATGATGT 11439
Qy 6242 GCATTAATGCCAGCGCGCT 6261
Db 11440 GCATTAATGCATCAGGCTGCT 11459

RESULT 13
AAV42364
ID AAV42364 standard; DNA; 16656 BP.
XX
AC AAV42364;
XX
DT 02-OCT-1998 (first entry)
XX
DE Representative eukaryotic layered vector initiation system sequence.
XX
KM Representative eukaryotic layered vector initiation system;
KM DNA alphavirus; structural protein expression; inhibit; pathogen;
OS immune response; stimulate; ss.
XX
OS Simble virus.
XX
FN US5789245-A.
XX
PD 04-AUG-1998.
XX
PF 30-OCT-1996; 96US-0741881.
XX
PR 15-MAR-1995; 95US-0404796.
PR 15-SEP-1993; 93US-0122791.
PR 18-FEB-1994; 94US-0198450.
PR 30-NOV-1994; 94US-0348472.
PR 20-JAN-1995; 95US-0376184.

OY 2320 GATACGAGATCATCTGCGAGGCTGATCATCATCTCTCTGTCACGTTAGCCGAAAGCG 2379
DB 7483 GGTATGAGGTAGACAAATATTACACCTGTCTCTACTGGCAATTGAAACTTTTGGCCAGAGA 7542
OY 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGCCCATCATCCTCTAGCGCTGACCTTAATAGG 2439
DB 7543 AAAGGCAATTCAGACCATCAAGAGGGAATTAACATCTCTACCGGTGTCTTAATAGT 7602
OY 2440 TGACGTAGTACACGACCTAC-----CCACGGGCGAATGTTCCATCC 2486
DB 7603 CAGCATAGTACTTTTCATCTGACTAATACACACACACACCTAGATTAGAGATTCT 7662
OY 2487 CTCAGCTGAATTTTCCACCAAGTTTACCTTACCAATCCGATGGCTTACCGAGATCCAAAC 2546
DB 7663 TTAACATCTGGGCGCGCGCCCTTCCGGGCCCACTGGCCATGTGGAGCCCGGAGAA 7722
OY 2547 CTCCTAGGCGCGCTGAGAGCGCTTTGGCCCCCGCTGGCTGTCTCAATTCGAAATCTTA 2606
DB 7723 GGAGGCAAGCGCGCCCGATGCTGCCCCGACGGGCTGGCTTCTCAATCCAGCAACTGA 7782
OY 2607 GGAGGTGATATGATCACTTGACTTTCAACAAACATCACCTTAATCCGCGCCAGGTCCA- 2665
DB 7783 CCACAGCGTCACTGCCCCCTTACTTGGACAGGCAACTAGACCTAACCCCACTGAC 7842
OY 2666 -----CCGCAAAAGAAAGAAAGAGTGTCTCTAAGCCAAACCTTACGCTTAAAGA 2720
DB 7843 GCCCGCCACCGCGCGAGAAAGAGAGCGGCCCAACCAACCAAGCCGAAAGCAACCA 7902
OY 2721 AGAGCAGCAAGCCAAAGAGAGCAAAAC--GCAGCTTAAACGAGGAAACGACAACTGA 2777
DB 7903 AAACCCAGAGAAAGAAAGCAACCTGCAAAACCCAAACCCGAAAGAGACGCGCA 7962
OY 2778 TGTGTATGAATGTGAGTCCGACAGACATTTCCATCATG---CTGAACGCGCCAGTGA 2834
DB 7963 TGGCACTTAAGTTGAGGCGCCGACAGATTGTTGACGTCAAGAACAGAGACGAGATGTCA 8022
OY 2835 ATGATATGCTCTGCTGTGCGAGGAGAGGCTGATGAACCACTCCAGCTTGAAGAA 2894
DB 8023 TCGGCGACGCACTGGCCATGAGAAAGATTAATAACTCTGACCTGAAAGAGAACCA 8082
OY 2895 TTGATATGAGCAATTGAGCGCGCTGAATTTGAAGAGGCTAGCATGTACGACTTGAAGT 2954
DB 8083 TCGACCACTGTGTCTATCAAAAGCTCAATTTTACCAAGTCTGACATATGACATGAGT 8142
OY 2955 AAGGGAAGTTCCTCCAGAACATGAATCAGACAGCTGACATACACGCAACAAACAC 3014
DB 8143 TCGCAGATTTGCGCACTCAACATGAGAGTGAAGCTTACCTTACACAGTGAACCCCG 8202
OY 3015 CGGCTTCTACCACTGGGACCCAGCGCGAGTCCAGTATGAGATGAGGATTTTACCTTAC 3074
DB 8203 AAGGATTTTATTAATCTGGACCCAGGAGCGGTGCAATATGAGAGGTATGATTTTACATCC 8262
OY 3075 CGAGAGAGTGGCGCGGAAAGGCGACAGCGAGAGACGATCTCTGACCAACAGAGCAGAG 3134
DB 8263 CTCGGGAGTGAAGAGCAGAGAGACAGCGGTCTGCGATCATGTATTAATCTCGGTGG 8322
OY 3135 TTGTGCTATTTGTTTGAAGAGTGCATATGAGGCGACGCTTACGCGCTTCACTGTGTCA 3194
DB 8323 TTGTGCTATTTGTTCTGAGTGGCGCTGATGAAGAGACGAACTTCTCGGTGCTCA 8382
OY 3195 CTGGAACCAAGAAAGGGGTGACATTAGGGAATACCCCGAAGGTTCTGACCGTGGT--- 3251
DB 8383 CCTGGAATATGTAAGAGAGACATTTAAGACGACCCCGAAGGAGCAGAAAGATGTCTCG 8442
OY 3252 -----CACTAGTTACAGCGCTATGCTGTTCGATGTCACTGTTCCATGCGCAAC 3305
DB 8443 CAGACACACTGTGTCACGGCAATGTGTTGTCTGGAATGTGAGCTTCCCATGCGACCGCC 8502
OY 3306 CACCGGTGTCTATTCATGACGCGCAAGAACCTGACGTGTCTGAAAGAAAGCTGCG 3365
DB 8503 CGCCCACTATGCTATACCGCGCAACTTCCAGAGCCTCGACATCTTTAAGAGAAAGCTGA 8562

OY 3366 ACAATCCAAATTACGACACGCTGTGAGAAAGTCTTGAATGTCCA---TCACCGCGGC 3422
DB 8563 ACCATGAGGCTTACGATACCTGTGCTCAATGCAATATTCGGGTGGGATCGTGTGCAAGA 8622
OY 3423 CCAAGCAGAGATTAACGATATCTTACACTGACCACTGACCTTACCTGGGTTCTGCCGT 3482
DB 8623 GCAAAAGAGGCTGTGTACACATTTTACCTGACCAAGCCCTTACTTGGGCACTGCTGT 8682
OY 3483 ATTGAGCACTCAACCGCGTGTTCAGCCCAATTAATTAAGAAAGTGTGGGAGCAAT 3542
DB 8683 ACTGCCACATATCTGAACCGGTGCTTACGCTTGTATGATGAGAGAGGTCTGGGAGCAAG 8742
OY 3543 CTGATGATGATGATTAAGATTCAGAGTCTGGGCAATTCGGGCTTACATGAGGAGCA 3602
DB 8743 CGAGATTAACACCATTAACGATACAGACTTCCGCCAGTTTGAATACGACCAAGCGGAG 8802
OY 3603 CTGGGATGTACCAAAATTCGTTTACATGTCTTTGACCAAGACCATATCAATCAAGAG 3662
DB 8803 CAGCAAGCGCAAAAGATACCGCTTACATGTGCTTAAGCAGGATACACCGTTAAAGAG 8862
OY 3663 ACAGTATGAGAAATAGCTATCAGCACTGACACCTGCGCTGTGGCCCAAG 3722
DB 8863 GACCATGATGACATCAAGATTAGCACCTGAGACGCTGTAAGAGCTTAGCTAGCAAG 8922
OY 3723 GGTACTTCTGTAGCTCAATGTCTCCAGAGTGAACGTATTAACGCTAGATTCAGAGCG 3782
DB 8923 GATACTTCTCTCTGCAAAATTCCTCCAGGGGACAGGTAAACGTTATGATAGATTA 8982
OY 3783 GAGCATCTGAGAAATTCATGACACCGTGAAGAAAGATCAGAGAAAGTTGTCTGTAGAG 3842
DB 8983 GCAACTCAGCAACCTCATGTACACTGCGCCCAAGATTAACAAATTCCTGGGACGGG 9042
OY 3843 AGAGTACTTGTTCACCCGCTGATGAAGCTGTAAATGCAACGTTTACATCACT 3902
DB 9043 AAAATATGATCTACTCTCCGTTTACCGTAAAGAAATTCCTTGAACATGTACACCTGTC 9102
OY 3903 TGAAGAGACGCTGCGCGGTATCAATCAATGACAGGCGCAGGCGCACCGCTTAAAGT 3962
DB 9103 TGAAG---ACAATCTGAGGCTACATCACTATGACAGGCGCGGACCGCATTTATCAT 9159
OY 3963 CTTATCTGAGAGAAAGCTCAGGCGAAGTTTACATTAACCACTTTGCGCAAGACGTCA 4022
DB 9160 CTTACTGGAAGATTCATCAAGGAAAGTTTACGAAAGCCGCACTTGGGAAAGCAATTA 9219
OY 4023 CCTACGATGTAATGTGCGGCACTACAGCAGAGTATGTATGACAGCGGAAAGAAATGA 4082
DB 9220 GGTATGATGTAAGTGGCGGCACTAACAGACCGGAACGTTTGAACCGGACCGAAATCA 9279
OY 4083 ACGGCTGCACTPAAAGCAAAACAGTGCATTGCTCAAGAGGAGCAACAAAGATGGTCT 4142
DB 9280 CTGATTGCAACCGCATCAAGAGATGCTGCTTATAGAGGAGCAACAAAGATGGTCT 9339
OY 4143 TCAACTGCGCGGATCTTATTAAGCAACAGACCACTCAGTGCAGATTAATTTGCACTTC 4202
DB 9340 TCAACTCAACCGGACTTGTATCAGATGACAGACCAAGCGGCAAGGAAATTTGCACTTTC 9399
OY 4203 CATTCGCTTACACCAAGTCTGCGCGGTTCCGTTTACGTTACACGCTTACAGTACGA 4262
DB 9400 CTTTCAAGTTGATCCCGGTGCTGCAATGTCTCTGTTGCGCAAGCGCGGATGTAATAC 9459
OY 4263 AGTGGTTCAAGGATCAACCTTCACTGATGCAATTCGACCAATTTGCTGACCAAG 4322
DB 9460 ATGCTTTAAACATATCACTGCTTCAATTTATGATACAGACACTTGACATTTGCTCACACCA 9519
OY 4323 GAAATTTGGGCTGCGAGACAGCAACAGCAAGATGATTAACAGGTCTTACATCCAGA 4382
DB 9520 GGAAGCTAGGGGCAAAACCGGAAACCAACCACTGATGATGATGCGAAAGACGCTGAGAA 9579
OY 4383 ATTTTCTGTGGGCGGAGAGGCTGAGATCTGATGGGTTAAACATTAACATCACTGAG 4442
DB 9580 ACTTCACGCTGACCGAGATGCGCTGGAATACATATGGGAAATCAATGACCACTGAGGG 9639
OY 4443 TCTGGGCCAGAGAGTCCGACCAAGGACCAACATGATGAGCGCATGATCATCATCC 4502

[illegible]

Db	10720	GGCTACTCAAGCCTTCGCCAANAAGATGCAATATGCCATACAGCGGCGGATACAGAT	10779
Qy	5583	ATGAATATGTGGAGAAACAATCAGACGACACCCCTGCAAGAAAACAGCAATTGGATGTA	5642
Db	10780	TTGAGATGTGGAAAAACAATCAGGCGGGCCACTGCAGGAAAAACCGACCTTCGCGGTGTA	10838
Qy	5643	AAATGTAAGTGGAGCCTCTGCGGCGCTTAACATGTGCTTAACGGGGCAATCCCTATCTCGA	5702
Db	10840	AGATGCAAGTAATCCGCTCCGCGCGGTGGACGTGTCAATCGGAAACATTTCCATTTCTTA	10898
Qy	5703	TTGACATCCCTGATGAGCTTTTGTGAGATCAATCAGAAATCACCMAAATTTTGAAGTTA	5762
Db	10900	TTGACATCCCGGAACGCTGCTTTATCAGGACATCAATGACACCACTGTCTCAACAGTCA	10955
Qy	5763	GCTGCACATGACGACATGCTATTATTCTGCAGACTTTGGTGTCTTCAACATTAACAGT	5822
Db	10960	AAATGTGAAGTCAAGTGTGACATTTATTCAAGACACTTCGCGGGATAGCCACCTGCACT	11019
Qy	5823	ACAAAGCTGAACAGGAGGGAACATGTTCAGTTCACTCCGACATCCACAGAGCTGTTTGA	5882
Db	11020	ATGATATCCGACCCGGAAGGTCAAATGCCGTAATTCGATTCGACAGACAGCAACTCTCC	11079
Qy	5883	AGGAAGGACCAACACATGTGACTGCCGTAGGCGAGCATAACTACTATTTTAGCACATCGA	5942
Db	11080	AAAGATGCACAGTACATGTCTCTGGAGAAAGAGCGGTGACAGTACACTTTAGCAACGGCA	11139
Qy	5943	GCCCAACAGCAAAATTTTATATGTTTGCCTATGCGGCAAGGCGAAGATCCACTGCAATGCTGAAT	6002
Db	11140	GTCACACAGCGGAACCTTATGTGTATCGCTGTGTGGAAAGAAACATCATGACAGAAAT	11189
Qy	6003	GTTAAACACCGCGCGACCACTAATTGGAGAAACCAATAAAGTCGACCAAGAAATCCAGG	6062
Db	11200	GTTAAACACCAAGTCGACCAATTCGTAGACACCCGCAACAAAATGACCAAGAAATTCAAAG	11258
Qy	6063	CGCGAGTTTCCAAAACATCTTGGAACTGCTGCTGCTTCACTGTGTGGGGAGAGATCATCCC	6122
Db	11260	CGCGACTCTCAAAAACATCATGGAAGTTGGCTGTTCCTTTTCGCGGGCGCTGCTGTC	11319
Qy	6123	TCAATTTGTAGGACTTTATAGTGTGTGCTGCAGCTCTATGCTTATAAACACAGTGAAT	6182
Db	11320	TATTAATTTATAGGACTTATGATTTTGTCTTGACAGCATGATGTGCTAGCACACGAAGAT	11379
Qy	6183	GACTGACCGGGGACACTGA-CATAGCGGTAAACCTGAGTACTTCCGAGGAAGCTGTGT	6241
Db	11380	GACCGCTACGCCCCCAATGATCCGACCAAGCAAACTGATGTACTTCCGAGGAAGTGAAT	11439
Qy	6242	GCATTAATGCCAGCGCGCGCT 6261	
Db	11440	GCATTAATGCATCAGGCTGAT 11459	
RESULT 14			
AAV70684			
ID	AAV70684	standard; DNA; 16656 BP.	
XX	AAV70684;		
AC			
XX	20-MAR-2003 (updated)		
DT	02-FEB-1999 (first entry)		
XX			
DE	Representative eukaryotic layered vector initiation system.		
XX			
KW	Eukaryotic layered vector initiation system;		
XX	alphavirus vector construct; gene therapy; ss.		
OS	Synthetic.		
XX	Sindbis virus.		
OS			
XX	US5843723-A.		
PN			
XX	01-DEC-1998.		
PD			
XX	30-OCT-1996; 96US-0739167.		
PF			

XX 15-MAR-1995; 95US-0404796.
PR 15-SEP-1993; 93US-0122791.
PR 18-FEB-1994; 94US-0198450.
PR 30-NOV-1994; 94US-0348472.
PR 20-JAN-1995; 95US-0376184.
XX
PA (CHIR) CHIRON CORP.
XX
PI Belli BA, Chang SMW, Driver DA, Dubensky TW, Ibanez CE;
PI Jolly DJ, Polo JM;
XX
DR WPI; 1999-044581/04.
XX
PT Alphavirus vectors constructs containing a 5' promoter of viral cDNA
PT by in vitro transcription - used in gene therapy
XX
PS Disclosure; Fig 3A-H; 140pp; English.
XX
CC The present sequence represents a representative eukaryotic layered
CC vector initiation system derived from Sindbis. The sequence exemplifies
CC the invention. The specification describes alphavirus vector constructs.
CC These constructs comprise a promoter 5' of viral cDNA which initiates
CC the synthesis of RNA from the viral cDNA by in vitro transcription,
CC followed by a 5' sequence which initiates transcription of alphavirus
CC RNA, followed by a nucleotide sequence encoding alphavirus nonstructural
CC proteins, a viral junction region which has been inactivated such that
CC viral transcription of a subgenomic fragment is prevented, an internal
CC ribosome entry site or a sequence which promotes ribosome read through
CC between adjacent reading frames, and an alphavirus RNA polymerase
CC recognition sequence. The recombinant alphavirus vectors can be used
CC for gene therapy.
CC (Updated on 20-MAR-2003 to correct PR field.)
XX
SQ Sequence 16656 BP; 4467 A; 4349 C; 4183 G; 3657 T; 0 other;
Query Match 36.5%; Score 2369.2; DB 20; Length 16656;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3693; Conservative 0; Mismatches 1983; Indels 44; Gaps 9;
QY 580 ACTGACGGTATGAAACGGGACGCTATATTTCTCATCGGAAACAGGCCAAGCTCCTTC 639
Db 5746 ACTACTACACCGGGGTAGGTGGGTACATATTTTCCAGCAGACAGGCCCTGGGCACTTGC 5805
QY 640 AACGAAATCAGTACGTACATATGTAACACAAAGCACTATTTGGATGGGCCCTCCATG 699
Db 5806 AAAAGAAAGTCCGTTCTGAGAACCAAGCTTACAGAACCCGACCTTGGAGGCAATGCTCGG 5865
QY 700 AGAAGTATTAAGCCCGCCGCTCGATCTCGAAAGAGAAATGTTACAGAAAGAACTGC 759
Db 5866 AAAGATTCATGCCCCCGGTGCTCGACAGGTGGAAGAGAACTCAAACTCAGGTACC 5925
QY 760 AATTATGCGCTCTGAAAGAAATAGAACAGGTATCAATCAAGAAATAGAAAATATGA 819
Db 5926 AGATGATGCCCAAGCAAGCAAAAGTAGTACAGTCTGTAAGTAGAATATCAGA 5985
QY 820 AAGCAATTAACGCGAGGAGATCTATTCTTGATTTGGGCAATATCTATCATCAGAAATGCA 879
Db 5886 AAGGCATTAACCACTGAGGAGTACTGTGAGACTACGACTGTATTAACCTCCACAG--- 6042
QY 880 ATTCCTGAGAGTGTATACAGATCAATTAATCTGTACCAATCTACTCGTCAACGGTATTA 939
Db 6403 ATCAGCCGAGATGCTATTAAGATCACTTATCCGAAACCACTTGTACTCCAGTACCTGCGG 6102
QY 940 ACAGGTTTACATCTGACAGAGTGGCGGTAAACGTGCACTTATAGTTATCCAGAGAAAT 999
Db 6103 CGAATCTACTCCGATCCACAGTTCGCTGTAGCTGTGTATCAACAATATCTGATGAGAACT 6162
QY 1000 ACCCTACAGTACGAGTATATGTATTAACAGATGAATACAGATGCGATCTTACATGATGCG 1059
Db 6163 ATTCGACAGTACATCTTATAGATTAAGTGAAGATGCTTAACTTGAATATGATGAG 6222
QY 1060 ACGGCGATCGTGTCTGTATAGATACAGCACTTTTGTCCGGCTTAACTGAGAAAGCTACC 1119

Db 6223 ACGGAGATGCGCCCTGCTGATATCTGCACTTCTGCCCCGCTTACGATTAAGATTACC 6282
QY 1120 CAAAGAAAGATAGCTATTTTTCAGCCAGAGATTAAGATCAGCCGTCCCATCCCTATACGA 1179
Db 6283 CGAAAAAATCATGAGTATTAAGACCCCGGATATCCGAGTGCAGTTCCATCAGCATGACGA 6342
QY 1180 ATACATTACAAATATGATTTGGCTGAGCTACTAAAGAAATGCAAGTTACCCAAATGCG 1239
Db 6343 ACAGCTACAAATATGCTCTATCTGCGTACCTTAAAGAAATTTGACATCAGCGATGCG 6402
QY 1240 GAGAAATTAACCTGTCTTATGATTTGGCGGATTTAAATGTTATGTTTCAAGAAATGCGAT 1299
Db 6403 GTGAACCTGCAACATGGAATCTGAGCAATTCATGATGATGATGCTTGAAGAAATATGAT 6462
QY 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
Db 6463 GTAATGACGATATTTGGAGAGATTTGCTCGGAAGCCAAATTAAGATTAACCACTGATTTG 6522
QY 1360 TTACGCAATATGACAAAGCTGAAGGCGGAAAGCGAGCAATTTGTTGCAATATCTC 1419
Db 6523 TCACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6582
QY 1420 ATAAATCTAAACCCGTTGAGAGATTAACATGATGATGATGATGATGATGATGATGATGATG 1479
Db 6583 ATAAATTTGCTCCATTTGAGAAAGATGCTATGATGATGATGATGATGATGATGATGATGATG 6642
QY 1480 ATGTCAAAATGATCTCCGCGACGAAACATACAGAGAGCGGCTTAAAGTGAAGTTATTC 1539
Db 6643 ACGTAAAGATTAACACAGGCGACGAAACACAGAAAGAAAGCCAAATTAAGTGAATAC 6702
QY 1540 AGGCTGACAGATCCCTGCTGATACGCTTACCTTGGGATGCGATGCGGAATTAAGCTGTA 1599
Db 6703 AAGCGGAGAACCCCTGGCGACATCTTATGATGATGATGATGATGATGATGATGATGATG 6762
QY 1600 GACTGAATGCGGCTCTTCTGCAAAATATCCATCTCTTTCGATGATGATGATGATGATGATG 1659
Db 6763 GCGTTAAGGCGGCTCTGCTGCTTCAAAACATTCACAGCTTTTGAATGATGATGATGATG 6822
QY 1660 TTGATGCGATATTTGCTGACATTTTCCACACGCGGACCCAGTATTTGGAACGACATCG 1719
Db 6823 TTGATGCAATATGACAGAACACATTCAGCAAGGCGACCCGCTGATGATGATGATGATG 6882
QY 1720 CGTGTGTTGATTAAGGAGAGAGACGCTATGCGCATTTGGCGTATGATGATGATGATGATG 1779
Db 6883 CATATTCGACAAAGCCAGACAGACGCTATGCGGTTAAACGCTGATGATGATGATGATG 6942
QY 1780 ACTTAGGTGACCAACCGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATG 1839
Db 6943 ACCTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7002
QY 1840 CTGTGACCTTACAGAAAGAGTTTAAATTTGCTGATGATGATGATGATGATGATGATGATG 1899
Db 7003 CCACCCATCTACACGGGTCTGTTTAAATTCGCGGCGATATGAAATCCGGAATGT 7062
QY 1900 TCTTAAGGCTGTTTGAACCACTAGTCAATATCATGATTTGCTTACGAGATTAACGATG 1959
Db 7063 TCTTCAACCTTTTGTCAACACAGTTTGAATGCTTATGCTTACGAGATTAACGATGATG 7122
QY 1960 AACGTTAAACCAAGCTGACGCGCGGCTCTATCGGCGAGATTAACATATGATGATGATG 2019
Db 7123 AGCGGCTTAAACGCTCAGATGTCAGAGCTTATGCGGAGACGACATCATATCATGAG 7182
QY 2020 TCGTCTCGACACCTTGAAGGAGAGATGCGCACTTGGCTGAACATGGAAGTAAAAA 2079
Db 7183 TAGATCTGACAAAGAAATGCTGAGAGTGCACCACTGCGCTCAACATGAGAGTTAAGA 7242
QY 2080 TTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2139
Db 7243 TCATCGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7302
QY 2140 ACCAGATTAACAGGACGACGCTGACAGATGCGAGACCTCTTAAAGGCTTTTAAAGCTTG 2199

D 7303 ATTGGTTACTTCACAGCGTGCCTGGCGGATCCCTGAAAAAGGCTGTTTAAGTTGG 7362
Q 2200 GAAACCATTTGGCAGTGTGATGATACCAACATCGCCGCCGCCCGGGGCACTGCATATG 2259
D 7363 GTAAACCGCTCCAGCGGACGACGAGCAAGACAGAAAGCGGCTCTGTGTATATG 7422
Q 2260 AACCAATGCGATGGAAACAGATTGGAATTAACGACGATGATGAAAGGCTGATATCCA 2319
D 7423 AAACAAAGGGGTGGTTTAAAGTAGATTAACAGGCACTTTAAGAGTGGCCGTGACGACC 7482
Q 2320 GATACGATGATATATGCGACAGGCTGATCATCATCTCTGTGTCACCTTAAGCCGAAGCG 2379
D 7483 GGTATAGGTGAGCAATATATTACCTGTCTACTGCGATTGAGAACTTTTGGCCAGAGCA 7542
Q 2380 TTAAAGACTTCAGAGGATTAAGAGGAGGCCCAATCACCTCTACGCTGACCTTAATAGG 2439
D 7543 AAAGAGCATTCAGACCATCATAGAGGGGAAATTAAGCATCTCTAGGAGTCTTAATAATG 7602
Q 2440 TGAAGTATGACACGCACTAC-----CCACCGGAGAAATGTTTCATACC 2486
D 7603 CAGCATGATATTTTCACTGACTAATATACAAACACACCATGAAATAGAGATTCT 7662
Q 2487 CTGAGTGAATCTTCCACAGTTTACCTTCAAAATCCGATGGCTTACCGAGATCAAAAC 2546
D 7663 TTAACATGCTCGGCGCGCCCTTCCGGCCCACTGCCATGTGAGGCGCGAGAA 7722
Q 2547 CTCTAGAGCCCGCTGAGGCGCTTGGGCCCGCGCTGGCTGCTCAATGGAATTTTA 2606
D 7723 GAGAGCGGCGGCGCCCATGCTGCGCAACGCGGCTGCTTCAATTCAGCACTGA 7782
Q 2607 GAGAGTGAATGATCACTTGAACCTTCAAAACAGATCACTTAATCCGCGCAGTCA- 2665
D 7783 CCAAGCGCTGCAAGGCCCTAGTATGTGAACAGGCAATAGACCTCAACCCCACTGTCAC 7842
Q 2666 -----CCGCAAAAGAGAGAGAGTCTCTTAAGCCAAACCTTCAAGCTTAATAA 2720
D 7843 GCCCGCACCGCGCCAGAAAGAGAGGCGCCCAAGCAACCAACCAACCGAAGAAACAA 7902
Q 2721 AGAAGCGAGAGCAAGCAAGAGCGAAAC---GCAAGCTTAACCGGGAACGCAAGCTA 2777
D 7903 AAACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7962
Q 2778 TGTGTATGAAGTGGAGTGGACAAAGACATTTCCGATCATG---CTGAACGCGCAATGTA 2834
D 7963 TGGCACTTAAGTGGAGGCGGACAGATGTTTTCGACGTCAGAAAGAGAGAGAGATGCA 8022
Q 2835 ATGAGATATGCTGCTGCTGTGGAGAGAGGCTGATGAACCACTCCAGCTTGAAGAGAAA 2894
D 8023 TCGGGCAGCAGCACTGGCATGAGAGAAAGTAATGAAACCTCTGCACTGTAAGAAACCA 8082
Q 2895 TTGATATGAGCAATTAGCGCGCTGAAATTGAAGAGGCTAGAGATGTAGGACTTGGAGT 2954
D 8083 TCGACCACTCTGTGCTATCAAAAGCTCAAAATTTACCAAGTGTGACATAGCAGATGAGT 8142
Q 2955 ACGGCGAGCTTCCCAAGACATGAATACAGACGCTGAGTACACAGAGCAAAACAC 3014
D 8143 TCGACAGTGTGCGATGACATGAGAGATGAGGATTCACCTTAACAAGTGAACACCCG 8202
Q 3015 CGGGCTTCTACACTGCGACCAAGCGGCACTGCTAGATATGAGATGAGATTAACCGTAC 3074
D 8203 AAGGATTTCTAATCTGGCACCAAGAGCGGCTGAGTATGAGAGTATGATTTCCATCC 8262
Q 3075 CGAGAGAGTGGGGGGAAGGCGACAGCGGAGACCGATCTCTGAGTATGAGTATGAGT 3134
D 8263 CTCGCGAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8322
Q 3135 TTGTGCTATATGTTCTAGAGAGTCAAAATGAGGCGAGCGTACGCGCTTCACTGATGCA 3194
D 8323 TTGTGCGGATGATCTCGGTGGCGCTGATGAGAGAACGAACTGCTTTCGGTGTGCA 8382
Q 3195 CTTGGAACAGAAAGGGGTGACATTAAGGATACCCCGAAAGGTTCTGAACCGTGT-- 3251
D 8383 CTGTGAATATGTAAGGAGAGCAATTTAAGAGCACTCCCGAAGGAGACAGAAAGATGTCCG 8442

Q 3252 -----CACTAGTTACAGCGCTATAGCGTCTTTCGAATGACAGTTCACATGCGAACAA 3305
D 8443 CAGACCACTGCTGACAGGCAATGTGTTGCTCGGAAATGTAGCTTCCATGCGACCGCC 8502
Q 3306 CACCGTGTGCTATTTACATGACCGCAAGAGAACTGCACTGCTGGAAGAGAGCTGCG 3365
D 8503 CGCCCAATGTATACCCCGGAACCTTCCAGAGCCCTCGACATCTTGAAGAGAAAGTGA 8562
Q 3366 ACAATCAATTTGACAGCTGCTGAGAGAGCTTTGAAATGTCCA---TCAAGCCGCG 3422
D 8563 ACCATGAGGCTTCGATACCTGTGCTCAATGCAATATGCGGTGCGGATGCTTGGCAGAA 8622
Q 3423 CCAAGAGCATTTACGATGATGATCTTCACTGACAGTACCTGCTGAGGCTTGCCTG 3482
D 8623 GCAAAAGAGAGCGCTGTTGAGCACTTTAACCCTGACAGCCCTTACTTGGGCAATGCTCT 8682
Q 3483 ATTGACAGCACTGACGCGCTGTTTCAAGCCCAATTAATTAAGAAAGTGTGGAGCGAT 3542
D 8683 ACTGCCACATGATGAACCGTGTCTGAGCCCTGTAAAGATCGAGAGGTCTGGGACGAA 8742
Q 3543 CTGATGATGATGATTAAGATTCAGAGTCTGCGCAATTCGGCTACATTCAGAGCGCA 3602
D 8743 CGAGCATTAACACATACGCTACAGATTCCTCCGCAATTTGATGATGACGAAAGCGAG 8802
Q 3603 CTGCGGATGTCACCAATTCGTTACATGCTTTGACACGACCAATGACATCAAGAGAG 3662
D 8803 CAGCAAGCGCAACAAAGTACCGCTACATGCTGCTTAAGAGATCAACCTGTTAAAGAG 8862
Q 3663 ACAATGAGAAATATGATTAAGACATCTGGAACCTGCGTCTTGTGSCACAAAG 3722
D 8863 GCACATGATGATGATCAATGATTAAGACCTCAGAGACGTTGTAAAGGCTTAAGCAAG 8922
Q 3723 GGTATCTCTGTTAGCTCAATGCTCTCCAGGTGACAGTGAACCTGATACAGAGCG 3782
D 8923 GATATCTTCTCTCGCAAAATGCTTCAGAGGAGACAGCTGAAGTGTAGATGTGATGA 8982
Q 3783 GACATCTGAGAAATATGATGACAGCTGAGAGAAAGATCAGAGAGAAATTTGTGCTGAG 3842
D 8983 GCAATCAGCAAGCTGATGATGACATGCGCCGCGAGATTAACCAAAATTCGTGGAGCGG 9042
Q 3843 AGGAGTACTTGTTCACACCGCTCATGAGAAAGCTGTGAAGTCCAGCTTACGATCACT 3902
D 9043 AAAAATATGATTAACCTCCGTTCAAGTAAAGAAATTCCTTGCACAGTGAACGCTC 9102
Q 3903 TGAAGAGAGCTGTGCGGCTGATTAACCAATGACAGGCGCAACGCGCTATTAAGT 3962
D 9103 TGAAG--ACAACTGCAAGGCTATCATCATATGACAGGCGGAGCCGACGCTTATATCAT 9159
Q 3963 CCAATCTGAGAGAGCGTCAAGGCGAAGTGTATTAACCAACCTTCTGCGAAGAGCTCA 4022
D 9160 CTAATCTGAGAGAGATCATGAGGAAAGTTTACCAAAAGCGCCATCTGGGAAAGACCTTA 9219
Q 4023 CTAAGATTAAGTGTGGAGTACAGCAGATATGTAAGACGCGAGCAAGAGATGA 4082
D 9220 GATATGATGATGAGTGGGAGCTTACAAAGCCGAAACCGTTGTGACCCGACGAAATCA 9279
Q 4083 ACGGCTGCAATTAAGCAAAACATGCAATGCTTCAAGAGGAGCAAAAGAAATGGGCT 4142
D 9280 CTGTGTGCAACCGCATCAACAGTGTGCTGCTATTAAGAGAGCAAAAGAGAGGCTCT 9339
Q 4143 TCAACTGCGGAGATTTATTAAGCAACAGACCACTCAGTGAAGTAAATTCACATTC 4202
D 9340 TCAATCAACCGAGCTTGTATGACATGACAGCAACACAGCGCCAAAGGAAATTCATTTGC 9399
Q 4203 CATTCGCTTGAACCGACAGTGTGCGCGTTCGTTAGCTCAACGCTTACAGTCAAG 4262
D 9400 CTTTCAAGTTGATCCCGGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 9459
Q 4263 AGTGTGTAAGAGGATACCTTCACTGACATGCAATGCAACCAATTTCTGACAGCA 4322
D 9460 ATGCTTTTAAACATCAGCTTCAATTAAGATCAGACCACTTGAATGCTCAACACCA 9519

QY 4323 GAAATTTGGGGCTGCGAGACAGCGCAAGCAAGAAATGATTAACAGGCTCTACATCCAGAGA 4382
Db 9520 GGAGACTAGGGGCAAAACCCGGAACCAACCACTGATGATGCTGGAAAGACGTCAGAA 9579
QY 4383 ATTTTCTGTGGGGGAGAAAGGGCTGAGTACGTATGGGGTACCATGAAACAGTCAAG 4442
Db 9580 ACTTCAACCGTGACAGAGATGGCTTGGAATACATATGGGAAATATGAGCCAGAGAGG 9639
QY 4443 TCTGGGCCAGAGTCCGAGACAGGCGACCCACATGATGAGCGCGCATGAGATCATCTCC 4502
Db 9640 TCTATGCCAAGAGTACAGCACAGAGACCTTCAAGATGGCCACACGAAATAGTACAC 9699
QY 4503 ACTATTTATCATCGGATCCAGTCTACACTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 4562
Db 9700 ATTACTACCATTCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9759
QY 4563 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4622
Db 9760 TGATGATTTGGGGTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9819
QY 4623 CGCCATAGCGGCTTGACCGAAGCGGATCCGACAGCATTTAGCGGTTTTGTGTGTGT 4682
Db 9820 CGCCATAGCGGCTTGACCGAAGCGGATCCGACATTTAGCGGTTTTGTGTGTGTGT 9879
QY 4683 TTGCGCCCAACCAACGCTGAAACATTTGGAGAAACCTTTGAACCATCTGTGTGTGT 4742
Db 9880 TTAGTGTGGCCCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9939
QY 4743 AACGTTTCTGTGGGACAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4802
Db 9940 AGCGTTTCTGTGGGACAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9999
QY 4803 GCTTTTCAATGCGATGCTTTTATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4862
Db 10000 ACTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10059
QY 4863 CTTTGGAAATGCGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4922
Db 10060 CTTTGGAAATGCGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10119
QY 4923 AACGCGAGGTTTACGCGCACTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4982
Db 10120 AAAAGGCGAGGTTTACGCGCACTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGT 10179
QY 4983 CTTTCAATGCGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5042
Db 10180 CTTTCAATGCGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10239
QY 5043 TTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5102
Db 10240 TCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10259
QY 5103 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5162
Db 10300 TCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10359
QY 5163 ACACAGCAATGAGAGGCGATACGTCGATTCGCTTCAAGCTGTGTGTGTGTGTGTGTGT 5222
Db 10360 ACACAGCAATGAGAGGCGATACGTCGATTCGCTTCAAGCTGTGTGTGTGTGTGTGT 10419
QY 5223 TCGCACTTAAAGTTTACACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5282
Db 10420 AGGGGATTAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10479
QY 5283 CCACCGCGCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5342
Db 10480 CTACCAAGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10539
QY 5343 AGGTATGCAAGGGCGGATATCAAGCGGCTTTTTCACCTTTTTCACCATTAAGTGTGTGT 5402
Db 10540 AAGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10599
QY 5403 GAAAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5462

Db 10600 ATCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10659
QY 5463 TCGGAGATATTCAGCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5522
Db 10660 TCGGAGATATTCAGCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10719
QY 5523 GAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5582
Db 10720 GAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10779
QY 5583 ATGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5642
Db 10780 TTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10839
QY 5643 AAATTTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5702
Db 10840 AGATTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10899
QY 5703 TTGACATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5762
Db 10900 TTGACATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10959
QY 5763 GCTGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5822
Db 10960 AATGTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11019
QY 5823 ACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5882
Db 11020 AATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11079
QY 5883 AGGAAGGAGCACACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5942
Db 11080 AAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11139
QY 5943 GCCCAAGCAAAATTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6002
Db 11140 GTCCACAGGCAACTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11159
QY 6003 GTAAACCAACCGGCGGACCACTTAATTGGAGAAACCATTAAGTGTGTGTGTGTGTGT 6062
Db 11200 GTAAACCAACCGGCGGACCACTTAATTGGAGAAACCATTAAGTGTGTGTGTGTGTGT 11259
QY 6063 CGGCAAGTTTCAAAACATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6122
Db 11260 CGGCAAGTTTCAAAACATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11319
QY 6123 TCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6182
Db 11320 TATTAATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11379
QY 6183 GACTGAGCGCGGACACATGA-CATNAGCGGTAAACCTGTGTGTGTGTGTGTGTGTGTGT 6241
Db 11380 GACCGCTACGCGCCCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11439
QY 6242 GCATTAATGCCAGGCGCGCT 6261
Db 11440 GCATTAATGCCAGGCGCGCT 11459

RESULT 15
AA292765
ID AA292765 standard; DNA; 16656 BP.
AA292765;
22-MAY-2000 (first entry)
Representative eukaryotic layered vector initiation system SEQ ID NO:1.
Recombinant alphavirus vector; gene therapy; anticancer; antiviral;
eukaryotic layered vector initiation system; antitubercular; cardiac;
antidiabetic; antineurodegeneration; immunomodulatory; immune response;
KW

OY 1960 AACGGTTAACCA CGTCAAGCTGCGCGCCCTCTATCGCGCAGCATATAATAGTCATGCTG 2019
DB 7123 AGCGGCTTAAAA CGTCCAGATGTGCAGGCGTTCA TTGGGAGACGA CAACATCATACATGAGG 7182
OY 2020 TCGTCTCCGACACCTTGATGCGGAGAGATGCGCACTTGGCTGAACATGAGTAAATA 2079
DB 7183 TAGTATCTGACAAAGAAATGGCTGAGAGGTGCGCACTGCGCTCAACATGAGGTTAAGA 7242
OY 2080 TTATGATGCAATTGTTGTTATCAAGCACTTACTCTGTGGGGGATTTATCCGTGG 2139
DB 7243 TCATGACGCGATCATCGGTGAGACACCTTACTCTGCGCGGATTTATCTTGCAAG 7302
OY 2140 ACCAGATTAACAGGCAAGCTGCGAGATGCGACAGCTCTTAAAGAGCTTTTAAGCTTG 2199
DB 7303 ATTGGTTACTTCCACAGCGTGC CGGTGGCGGATCCCTGAAAGGCGCTTTAAGTTGG 7362
OY 2200 GAAACCACTTGGCAATCGATGATACC CAAAGCTGCGACCGCGCGGCACTGCATGATG 2259
DB 7363 GTAAACCGCTCCACCGCGACGACGAGCAAGACGAAGACGCGCTGTGCTAGATG 7422
OY 2260 AAGCAATGCGATGGAACGAATTGGAATTAGCGAGATGATGAAAGCGCGTGAATCCA 2319
DB 7423 AAACAAAGCGGTGTTTAGAGTAGTATPAACAGGCACTTTAGCAGTGGCGTGACGACCC 7482
OY 2320 GATACGAGATCATATCGCGAGGCTGATCATCAGCTCTGTGCAAGTTAGCGGAAAGCG 2379
DB 7483 GGTATGAGTAGACAAATTTACACTGTCTACTGGCATTTGAGAACTTTGGCCAGAGCA 7542
OY 2380 TTAAAGACTTCAGAGCATPAAGAGGAGCCCAATCACTCTTACGCGCTGACCTAAATAG 2439
DB 7543 AAAGAGCAATCCCAACCCATCAGAGGGGAAATTAACATCTCTACGCTGCTCTTAATAGT 7602
OY 2440 TGAGTAGTAGACAGCACTTAC-----CCACCGGCGAGATGTTCCATACC 2486
DB 7603 CAGCATAGTACATTTCACTGTGCTAATCTACACACACACACTGATAGAGATCT 7662
OY 2487 CTCAGCTGAATTTTCCACCAAGTTTACCTTACCAATCCGATGGCTTACCGAGATCCAAAC 2546
DB 7663 TTAACTGTCTGGCGCGCGCCCTTCCGCGCCCACTGCCATGTGAGAGCGCGGAGAA 7722
OY 2547 CTCTAGGCGCGCTGAGAGCGGCTTTCCGCGCCGCGCTGCTGCTCAATTCGAAGATCTTA 2606
DB 7723 GAGAGCAGCGCGCCCGCATGTCTGCCCAACGCGCTGGCTTCTCAAACTCAGCACTGA 7782
OY 2607 GGAAGTCATAGTCAATTTGAATTTCAACACAGATCACCTAATCCGCGCCAGGTCCA- 2665
DB 7783 CCACAGCCGTCAAGTCCCTTCACTTGAACAGGCAATGACCTCAACCCCACTGCAC 7842
OY 2666 -----CCGCCAAAGAAAGAAAGTGTCTCTAAGCCAAACCTTACTGAGCTTAAAGA 2720
DB 7843 GCCCGCCACCGCGCCAGAAAGAGAGCGCGCCCAACCAACCGAAGCCGAAAGAAACCA 7902
OY 2721 AGAAGCAGCAAGCCAAAGAGAGCAAA--GCAGCTTAAACGAGGAAACGACAACTGA 2777
DB 7903 AAACGCAAGAGAAAGAAAGAAAGCAACTGCAAAACCCAAACCCGAAAGAGAGCGCA 7962
OY 2778 TGTGTATGAAGTTGAGAGTCGACAGACATTTCCATGATG---CTGAAGCGCCAAAGTGA 2834
DB 7963 TGGCATTAAGTTGAGAGCGCGCAAGATTGTCAGCTCAAGAACGAGAGCGAGATGTCA 8022
OY 2835 ATGATATGCTGCTGCTTTCGAGAGAGAGCTGATGAACCACTCCAGCTGAAAGGAAATA 2894
DB 8023 TCGGGCAGGCACTGCGCATGAGAAAGAGTAAATTAACCTTGCACGTGAAGAGAACCA 8082
OY 2895 TTGATATGAGCAATTGACGCGCGCTGAAATTGAAGAGGCTAGCATGTGACGCTTGAGT 2954
DB 8083 TCGACCACTGTGTCTATCAAAAGCTCAATTTAACAGATGCTCAGCATATACGATGAGT 8142
OY 2955 ACGGGAAGCTTCCCGAAGCATGAATCAGACAGCTGACATGACACGAGCAAAACAC 3014
DB 8143 TCGCACAAGTTGCGACTCAACATGAGAGGACTTCACTTACACAGATGAACACCCG 8202

OY 3015 CGGCTTCTACAACTGGACCAACGCGGAGCTCCAGTATGAGAAATGAGATTTACCGTAC 3074
DB 8203 AAGATTTCTATTAATCGGACCAACGAGCGGTGACATTAATGAGAGGTGATTTACATCC 8262
OY 3075 CGAGAGATGTTGGCGGGAAGGCGACAGCGGAAACCGATCTCTGACAAACAGAGCAGAG 3134
DB 8263 CTCGCGAGATGAGAGGACAGAGAGACAGCGGCTGCTCCATCATGTGATTAATCCGCTCGG 8322
OY 3145 TTGTGCGATTTGTTCTAGAGAGGTGAATGAGGGGACCGGTACCGGCTTTCAAGTGTCA 3194
DB 8323 TTGTGCGATTTGTTCTGAGGTGCTGATGAAAGAAACGAATCTCCCTTGTGCTGCA 8382
OY 3195 CTTGGAACAGAAAGGGGTGACCATTAAGGATATCCCGGAAGTTCTGAACCGTGT--- 3251
DB 8383 CTTGGAATTAATGAAAGGAAGCAATTAAGAGACCCCGGAAGGACAGAAAGTGTGCG 8442
OY 3252 -----CACTAGTTACAGCGCTAGTGTGCTTTGCAATGTCACTTTCCATGCAAAAC 3305
DB 8443 CAGCACACTGTGTACCGGCAATGTGTGTCTCGAAATGTGAGCTTCCATGCGACCGCC 8502
OY 3306 CACCGGTGTGCTATTCATGACCGGCAACGAAACGACCTGAGGTGCTGAAAGAAAGTGG 3365
DB 8503 CGCCCAATGCTATACCCCGCAACCTTCAAGAGCCCTGACATCTTGAAGAAAGTGA 8562
OY 3366 ACAATCCAAATTTACGACACAGCTGCTGAGAAAGTCTTGAATGTCCA---TCACCGCGGC 3422
DB 8563 AOCATGAGGCGCTAGCATACCTGTCTCATGTGCATTTGCGGTGCGGATCTGTGCGAAGA 8622
OY 3423 CCAAAAGAAATTAACGATGACCTTACATGACATGTCCTTACTGCGGCTTCTGCGCT 3482
DB 8623 GCAAAAGAAAGGCTGCTTACCATTTTACCTGACAGAGCCCTTACTTGGGCACTGCTGT 8682
OY 3483 APTGAGACATCAACGCGGTGTTTCAAGCCCAATTAATGAGAACGTGCGGAGCAAT 3542
DB 8683 ACTGCACATATCTGAACCGGTCTTCAAGCCCTGTTAATGATGAGAGGTCTGCGACGAAG 8742
OY 3543 CTGATGATGATTCGATTAATGATTCAGGTCTGCGGCAATTTGCGCTACATCAGGAGCA 3602
DB 8743 CGGAGATTAACACATTAAGCATACGATTCGCGCCCACTTTGAGATGACACAAACGAG 8802
OY 3603 CTGCGGATGTACCAATTTCCGTTATCATGTTCTTTCGACCAAGACCATATCAAGAGAG 3662
DB 8803 CAGCAAGCGCAAAACGATACCGCTACATGCTGCTTAAACAGATACACCCGTTAAAGAG 8862
OY 3663 ACAGATGAGAAATTAATCTATCAGCAGATCTGAGCCCTGCTGTCTTGGCCCAAG 3722
DB 8863 GCACCATGATGACATCAAGATTTAGCACCTTGAAGCGGTGAGAGGCTTAGCTTAAAG 8922
OY 3723 GGTACTTCTGTAGCTCAATGTCTCTCAGGTGACAGTGAACGCTCAGTATCACAGAGC 3782
DB 8923 GATACTTCTCTCTGCAAAATGCTCTTCCAGGGGACAGGCTTAAAGCTATGATGATGA 8982
OY 3783 GAGCATCTGAGAAATTCATGACCGGTGAGAAAGATCAGAGAAAGTTGTGCTAGAG 3842
DB 8983 GCAACTCAGCAACGTCATGATACATGCTGCCCAAGATTAACAAATTTCCGCGGAGCGG 9042
OY 3843 AGAGTACTTGTTCACACCGGCTCATGAGAAAGCTGTAAAGTCCACGTTTACATCACT 3902
DB 9043 AAAATATGATCTACTCTCCGTTCAAGGTAAAGAAATTTCTTTCACAGTATGACCGCTC 9102
OY 3903 TGAAGGAGAGCTGCGCGGTATACATTAACATGACAGGCGCCAGGCGCTGATTAAGT 3962
DB 9103 TGAAG---ACAATGCGAGGCTATCACTATGACAGGCGCGGACCGCATGATTAAC 9159
OY 3963 CTTATCTGAGAAAGGCTCAGGCGAAGTGTACATTTAAACCACTTTCTGCAAGAAAGCTCA 4022
DB 9160 CTTACTCTGAGAAATCATCAGGAAAGTTTACGCAAAACCCGCACTCTGGAAGAAACATTA 9219
OY 4023 CTTACGAATTTAATGTGCGCATACAGCACAAGATATGCTGAGACGCGGAACGAAGATGA 4082
DB 9220 GGTATGAGTGAAGTGGCGGACATTAAGAGACCGGAACGTTTTCGACCGGCAACGAAATCA 9279
OY 4083 ACGGCTGACATAAGCAAAACAGTGCATTTGCTTCAAGAGGAGCAACAAAGAAATGCTCT 4142

Db 9280 CTGGTTGACACCGCCATCAAGAGTGCCTCGCTTATTAAGAGGACCAACGAAAGGGCTCT 9339
Qy 4143 TCAACTCGCGGATCTTATTAAGGCAACAGACCACTGATGCAAGTAATTGCACTTC 4202
Db 9340 TCAACTCACCGGACTTATGATGACATGACGACCAACGCGCCCAAGGAAATTTGATTTGC 9339
Qy 4203 CATTCGCTTGAACCGACAGTGCCTGGGTTCCGTTAGTGCACACGCTTACAGTCAAG 4262
Db 9400 CTTTCAAGTTGATCCCGGTCCTGCAATGCTCTGTTGCCACGCGCCGAAATGTAATAC 9459
Qy 4263 AGTGGTTCAAAAGGATCAACCTCCACCTGACTGCAATGCGACCAATGCTGCAACGA 4322
Db 9460 ATGGCTTTAAACATCATGAGCTCCAAATTAGTATACAGACCACTTGACATTTGTCACCA 9519
Qy 4323 GAAATTTGGGCTGCGACGACGCAACAGCAATGATTAACAGGCTTATCCTCAAG 4382
Db 9520 GGAGACTAGGGGGGCAACCCCGAAACCAACCACTGAATGATCTGTGAAAGACGCTCA 9579
Qy 4383 ATTTTCTGTGGGGGCGAAGGGCTGAGTACGTATGGGGTAAACATGAACAGTCAAG 4442
Db 9580 ACTTACCGTGCACGAGATGAGCTGGAATACATATGAGGAATCATGAGCAGATGAGG 9639
Qy 4443 TCTGGGCGCAGAGTCCGCAACGACGCAACCACTGATGCGCATGATCATCATCC 4502
Db 9640 TCTATGCCAAGAGTCAACGACGAGAGACCTCTCAGATGCGCCACAGAAATAGTACAGC 9639
Qy 4503 ACTATTAATCATCGGATCTCAGTCTACACTGTCAATTGCTGTGTGCTGCTTCTCTA 4562
Db 9700 ATTACTACATCGCATCGCTCTGTGTACCACTTAAAGCGTGCATCAGCTACCGTGC 9759
Qy 4563 TCCGTGAGGCACTGCATCATCAGCAGCTTGCAATGCGCAAGCAAGAAAGACGCTG 4622
Db 9760 TGAATGTGCGTAACTGTTCAGTGTGATGCTGTAAGCCGCGGAGTGCCTGA 9819
Qy 4623 CGCATATCGGCTTGCAACGCAACGCAACGATACCCACAGCATTAAGCGGTTTGTGCTGA 4682
Db 9820 CGCATATCGGCTTGCAACGCAACGCGGTAATCCCACTGCGTGCCTTGTGCTGCG 9879
Qy 4683 TTGGGCCAACCAACGCTGAAACATTTGAGAAATTTGAACCATCTGTGTTAAACA 4742
Db 9880 TTAGTGTGCGCAATGTGAAACGTTCAACGAGACATGAGTACTTGTGTGCGAAGATC 9939
Qy 4743 AACGTTCTCTGGGACAGTGTGCACTCTCTGGAAGCGCTGTAATCTGTTCCGCT 4802
Db 9940 AGCGTTCTCTGGGACAGTGTGCACTTGTGCGGCGCTTATGTTCTAAATGCGCT 9999
Qy 4803 GCTTTTCATGCTGATGCTTTTATTTGTTGAGGCGTCTGCTGCGGGAAGGTAGACG 4862
Db 10000 ACTGCTCTGCTGCTGCTTTTATGTTGTTGCGGCGCTTACTGCGGAAGTACG 10059
Qy 4863 CTTTCAACATGCGACCACTGTGCAATGTTCCGGGATCCGCTATAGGCGTTGCTG 4922
Db 10060 CTTAAGCAATGCGACCACTGTGCAATGTTCCAAATGTGCCAAGATACCTTAAGGCACTTGTG 10119
Qy 4923 AAGCGAGGTTACGGGCACTTAACCTGAGATCAGGTCGTCACTGGAATTAACAC 4982
Db 10120 AAGGGAGGATGAGCCCGCTCAATTTGAGATCATGTAATCTCTCGAGGTTTTC 10179
Qy 4983 CTTTCACTAACAGAGATGATGACTGCAATTCACACAGTCAATCTTTCACACAG 5042
Db 10180 CTTTCAACCAACAGAGATGATGACTGCAATTTACACACAGTGTGCTCCCTCCCAAAA 10239
Qy 5043 TTAATGCTGCGGCTCTCGAGTGAAGGATCTTCAAGGCGGATTAACATGCGCG 5102
Db 10240 TCAATGCTGCGGCTCTCGAGATGTCAGCGCGCTCATGACATTAATCTGCAAG 10299
Qy 5103 TTTTGGCGGTGTACTCTTTCAATGCGGAGCGCAATGCTTCTGACAGTGA 5162
Db 10300 TCTTGGAGGGGTCTACCCCTTAATGTGGGAGAGGCAATGTTTGTGCAAGTGA 10359
Qy 5163 ACAACAATGATGAGGCGTACGTGAGTTCCGTCAGACTGATATGATCAACGAG 5222

Db 10360 ACAACGATGATGAGGCGCTGATCGAATGTCTAGCAGATTGCGCTGACCAACGCG 10419
Qy 5223 TCGCATTAAGTTTCAACAGCTGTCTGAAATGCGGCTCGCTATATGTTAGGGACA 5282
Db 10420 AGCGATTAAGGTGCAACACTGCGCGATGAAGATGAGACTGCTATATGTAGGGACA 10479
Qy 5283 CCAACGCGACCTGATACCTTTGTCAATGGCTACAGCGAGGTCTCTCAGGGACCTGA 5342
Db 10480 CTACAGTTTCTTATGATGTGTAGTGAACGAGTCAACACGAGAACCTTAAAGCTTGA 10539
Qy 5343 AGTCAATGACGAGCGGATATCAAGCGCTTTTCACTTTGACATAGGTGTCTATCA 5402
Db 10540 AAGTCAATGAGCAACCAATTTCAAGATCGTTTACGCTATGATCAATAGGTGTATCC 10599
Qy 5403 GAAAGGCGCTTTTCAATGACATTCCTCGATGATGAGGATGAAACGAGAGCT 5462
Db 10600 ATCGCGCTGTGGTGAACATTAAGTCTTCCGAAATGAGAGGATGAACGAGAGCT 10659
Qy 5463 TCGCGATATTCAAGCATCTCGCTTGATGCTACAGATAGTAGCCGCACTGACATAC 5522
Db 10660 TCGGAGCATTTCAAGTACTCTCTGACTAGCAAGATCTCATGCGCAGACAGACTTA 10719
Qy 5523 GCGTCTGAAGCTTCTGTCAAGAAATCCAGTCCCTTACACCAAGCATCAAGGT 5582
Db 10720 GCGTACTCAAGCTTCCGCAAGAAAGTGCATGTCCGTACAGGACGCGCATCAAGAT 10779
Qy 5583 ATGAATGTGGAAGAACATCTCAGACGACCCCTGCAAGAAACAGCACTTTGATGTA 5642
Db 10780 TTGAGATGTGGAAGAACATCTCAGGCGCCCTGCAAGAAACCGCACTTTGGGGTGA 10839
Qy 5643 AAATGGAATGAGAGCTCTGCGAGCGCTTAAGTGTGTTAGGGACATCTTATCTGA 5702
Db 10840 AGATGAGTAATTCGCTCCGAGCGGTGAGCTGTTCATACGGGAACTTCCATTTCTA 10899
Qy 5703 TTGACATCTCTGATGACGCTTTTGTGAGATCATAGATCAACCAATTTTGAAGTGA 5762
Db 10900 TTGACATCTCTGATGACGCTTTTATGAGACATCAGATGACCACTGTCTCAACAGTCA 10959
Qy 5763 GCTGCACATGACGAGTGTATTTTGTGACACTTTGTGTGTTCTTAACATTAAGT 5822
Db 10960 AATGTGAATGATGATGATGATTAATTCAGCAACTTGGGAGTGCACCTGAGT 11019
Qy 5823 ACAAGCTGACAGGAGGAGCATTTGTCAGATTCACATCCACCTCAGACAGCTGTTTGA 5882
Db 11020 ATGTATTCGACCGGAGAGTCAATGCCCTGATTTGCAATTTGAGACAGCAACTTCC 11079
Qy 5883 AGAAGGACCAACATGATGATGCTGAGGACAGCATTAACATTTTGAACATGCA 5942
Db 11080 AAGATGCAAGTAACTATGCTCGAGAAAGAGACGGTGACAGTAACTTTTGAACCGCA 11139
Qy 5943 GCCCACAAGAAATTTTATGTTTCTGATGCGGCAAGATCCACCTGCAATGCTGAAT 6002
Db 11140 GTCCACAGGGAATTTATGTTATGCTGTGTGTGGAAGAAACAACTGCAATGCAAT 11199
Qy 6003 GTAAACCAACGCGGACCAATATTTGAGAAACCACTAAAGTGCACCAAGAAATTCAG 6062
Db 11200 GTAAACCAACGAGTGAACATATGATGAGACCCCGCACAAAATGACCAAGAAATTTCAAG 11259
Qy 6063 CGGAGTTTCAAAACATCTTGAATGAGTCTGCTGCTGCTGCTGTTGGGAGATCATCC 6122
Db 11260 CCGCATCTCAAAAACATCATGAGATGAGTGTGCTGCTTTCGCGGCGCTCGCTGCG 11319
Qy 6123 TCAATGTTGAGACTTATATGTTGTGTGAGCTCTATAGCTTATTAACACAGTAT 6182
Db 11320 TATTAATTAATGAGACTTATATTTTGTGAGATGAGTGTGATGACACAGAAAT 11379
Qy 6183 GACTGACGCGGACGACTGA-CATAGCGTAAACTGATGATCTTCCAGAAAGCTGTGT 6241
Db 11380 GACGCTACGCGCCCAATGATCCGACAGCAAACTGATATCTTCCGAGAACTGATGT 11439
Qy 6242 GCATTAATGCAACGCGCGCT 6261
Db 11440 GCATTAATGCAACGCGCTGT 11459

Mon Nov 17 10:39:24 2003

us-10-023-649-1_copy_5000_11484.rng

Page 49

Search completed: November 15, 2003, 23:18:21
Job time : 1555 secs

Db 5060 CGGCGATTTCTAGAGACACCAAGTGGGATCACTGCTATGCTTGGGGGGCTCTGTTCT 5119
Qy 121 CTAGCTCTGCCACACGCTCGTCAACCGCTGGAGCGACTATGACAGCGGTTTGTGTTA 180
Db 5120 CTAGCTCTGCCACACGCTCGTCAACCGCTGGAGCGACTATGACAGCGGTTTGTGTTA 5179
Qy 181 CAGCTGATGTCATCAAGCGAACGCTCTAGTGGAGATCCCTAGTGCCTGGCTTGG 240
Db 5180 CAGCTGATGTCATCAAGCGAACGCTCTAGTGGAGATCCCTAGTGCCTGGCTTGG 5239
Qy 241 ACGTCCAGCTGCTTCTGACGTCATGATTCACCTGAGATTTCCAAATGTCATCAAGCT 300
Db 5240 ACGTCCAGCTGCTTCTGACGTCATGATTCACCTGAGATTTCCAAATGTCATCAAGCT 5239
Qy 301 TTGAAGTGAACACCATCTGTACAGACCTTAATCTGGGAGTGTGGAGCTGTGGAC 360
Db 5300 TTGAAGTGAACACCATCTGTACAGACCTTAATCTGGGAGTGTGGAGCTGTGGAC 5359
Qy 361 TGGCGGAAATTAATGACGACTTCAATCTGCTCCCTTCGAGTTCTTTCGAGCTACAGAC 420
Db 5360 TGGCGGAAATTAATGACGACTTCAATCTGCTCCCTTCGAGTTCTTTCGAGCTACAGAC 5419
Qy 421 CAGTACCGGACACACGAGACGCCCCATCCCATCAGTACGACGCTTCCGACCTC 480
Db 5420 CAGTACCGGACACACGAGACGCCCCATCCCATCAGTACGACGCTTCCGACCTC 5479
Qy 481 CAGTTCGAAAGCCACGACGACCTAACTCAACACCAACCAAGAGTGCCTTGAAGCATCT 540
Db 5480 CAGTTCGAAAGCCACGACGACCTAACTCAACACCAACCAAGAGTGCCTTGAAGCATCT 5539
Qy 541 CAGAGCGGAGTGAAGAGATCATCGGTCAACCTCAATGAGAGTGAAGAGAGAG 600
Db 5540 CAGAGCGGAGTGAAGAGATCATCGGTCAACCTCAATGAGAGTGAAGAGAGAG 5539
Qy 601 CGTATATTTTCTCATCGGAAACAGGCAAGGTCACTTTCACAGAAATCACTAGTCAAT 660
Db 5600 CGTATATTTTCTCATCGGAAACAGGCAAGGTCACTTTCACAGAAATCACTAGTCAAT 5659
Qy 661 GTAAACCTACAGAAACCTATATTGGATCGGGCCGTCCATGAGAGATTTACCCCGCGCC 720
Db 5660 GTAAACCTACAGAAACCTATATTGGATCGGGCCGTCCATGAGAGATTTACCCCGCGCC 5719
Qy 721 TCGATCTCGAAGAGAGAGAAATGTTACAGAGAACTGCAATTTGCGGCTCTGAGGAA 780
Db 5720 TCGATCTCGAAGAGAGAGAAATGTTACAGAGAACTGCAATTTGCGGCTCTGAGGAA 5779
Qy 781 ATAGAAGAGGTATCAATCAAGAAAGTGAATAATGAAAGCAATTAACAGCGAGCGAC 840
Db 5780 ATAGAAGAGGTATCAATCAAGAAAGTGAATAATGAAAGCAATTAACAGCGAGCGAC 5839
Qy 841 TCATTTCTGGATTTGGGACCATATCTATCATCAAGAGTGAATCCGTGAGTGTAAACAGAG 900
Db 5840 TCATTTCTGGATTTGGGACCATATCTATCATCAAGAGTGAATCCGTGAGTGTAAACAGAG 5899
Qy 901 TCAATTTCCGTACCAATCTACTCGTCAAGGTAAATTAACAGGTTTACATCTGAGAGG 960
Db 5900 TCAATTTCCGTACCAATCTACTCGTCAAGGTAAATTAACAGGTTTACATCTGAGAGG 5959
Qy 961 TCGCGGTTAAACGTGCAACTTAATTATCAAGAAATTAACCTTACAGTACGAGTTATTT 1020
Db 5960 TCGCGGTTAAACGTGCAACTTAATTATCAAGAAATTAACCTTACAGTACGAGTTATTT 6019
Qy 1021 GTATTAACAGATGAATACGATCGTATCTTGAATGATGTCAGCGGCTGCTGTCTAG 1080
Db 6020 GTATTAACAGATGAATACGATCGTATCTTGAATGATGTCAGCGGCTGCTGTCTAG 6079
Qy 1081 ATACAGCCACTTTTGTCCGGCTAAACGAGAGTACCCAAAGAGATGCTATTTGCT 1140
Db 6080 ATACAGCCACTTTTGTCCGGCTAAACGAGAGTACCCAAAGAGATGCTATTTGCT 6139
Qy 1141 AGCCAGAGATTAAGATCAAGCGCTCCCATGCTTATCAGAAATTAATTAATGATTTGG 1200

Db 6140 AGCCAGAGATTAAGATCAAGCGCTCCCATGCTTATCAGAAATTAATTAATGATTTGG 6199
Qy 1201 CTGACGTACTTAAGAAATTTGCAAGTTAACCAATGCGAGAAATTAACCTGTATGATT 1260
Db 6200 CTGACGTACTTAAGAAATTTGCAAGTTAACCAATGCGAGAAATTAACCTGTATGATT 6259
Qy 1261 CGCGGCAATTAATGTTGATTTGTTCAAGAAATATCAGATGATGATGATCTGGGATA 1320
Db 6260 CGCGGCAATTAATGTTGATTTGTTCAAGAAATATCAGATGATGATGATCTGGGATA 6319
Qy 1321 CTTTGGCGATTAACCTTATTTGGCTTAATCTACAGAGAACTTACGAAATATGTCAGAAC 1380
Db 6320 CTTTGGCGATTAACCTTATTTGGCTTAATCTACAGAGAACTTACGAAATATGTCAGAAC 6379
Qy 1381 TGAAGGGCCGAAAGACGACGATGTTGTCGAATCTCATATCTTAAACGTTGAGG 1440
Db 6380 TGAAGGGCCGAAAGACGACGATGTTGTCGAATCTCATATCTTAAACGTTGAGG 6439
Qy 1441 AGATACCAATGATCAATTTGTCATGATCTTAAAGAGAGATGTCAAGTTACTCCGGCA 1500
Db 6440 AGATACCAATGATCAATTTGTCATGATCTTAAAGAGAGATGTCAAGTTACTCCGGCA 6499
Qy 1501 CGAAACATACAGAGAGACGCGCTTAAGGTGAGTTATTCAGGCTGCAAGATCCCTTGTCTA 1560
Db 6500 CGAAACATACAGAGAGACGCGCTTAAGGTGAGTTATTCAGGCTGCAAGATCCCTTGTCTA 6559
Qy 1561 CGGCTTAACCTTGGCGGATTCATCGGGAATTAATCGGTGAATGAGTGGGCTTCTGC 1620
Db 6560 CGGCTTAACCTTGGCGGATTCATCGGGAATTAATCGGTGAATGAGTGGGCTTCTGC 6619
Qy 1621 CAAATATCACTACTCTTTCGACATGTCAGCGAGATTTTGTGATGATTTGCTGAAC 1680
Db 6620 CAAATATCACTACTCTTTCGACATGTCAGCGAGATTTTGTGATGATTTGCTGAAC 6679
Qy 1681 ATTTCCACACGCGGACCCAGTATTTGAAACGACATCGGCTGTTGATTAAGAGGAG 1740
Db 6680 ATTTCCACACGCGGACCCAGTATTTGAAACGACATCGGCTGTTGATTAAGAGGAG 6739
Qy 1741 ACGACGCTATCGCATTTGGCGGTTGATGATCTTGAAGGCTTAAGGTGTGACCAACCGC 1800
Db 6740 ACGACGCTATCGCATTTGGCGGTTGATGATCTTGAAGGCTTAAGGTGTGACCAACCGC 6799
Qy 1801 TCTTAGATTGATAGAGCGCGCTTCGCAATATCACTGTCACCTTACCTACAGAA 1860
Db 6800 TCTTAGATTGATAGAGCGCGCTTCGCAATATCACTGTCACCTTACCTACAGAA 6859
Qy 1861 CGAGTTTAAATTTGGTGCATGATGAATCCGATATGTTCTTAAACGCTGTTGTCAACA 1920
Db 6860 CGAGTTTAAATTTGGTGCATGATGAATCCGATATGTTCTTAAACGCTGTTGTCAACA 6919
Qy 1921 CACTAGCAATATCAATGATTTGTCAGAGATCTAGCTGAACGGTTAACACGTCAGCGCT 1980
Db 6920 CACTAGCAATATCAATGATTTGTCAGAGATCTAGCTGAACGGTTAACACGTCAGCGCT 6979
Qy 1981 GCGGCGCTCTATCGGCGACGATTAACATAGTGAATGTCGTCTCGACACCTTGAATG 2040
Db 6980 GCGGCGCTCTATCGGCGACGATTAACATAGTGAATGTCGTCTCGACACCTTGAATG 7039
Qy 2041 CGAGAGATGCGCATCTTGGCTGAACATGGAAGTAAATTAATGATGCAATTTGTTA 2100
Db 7040 CGAGAGATGCGCATCTTGGCTGAACATGGAAGTAAATTAATGATGCAATTTGTTA 7099
Qy 2101 TCAAGAGACCTTACTTGTGGGGGATTTATCCTGCTGAGACAGATTAACAGGACAGCGCT 2160
Db 7100 TCAAGAGACCTTACTTGTGGGGGATTTATCCTGCTGAGACAGATTAACAGGACAGCGCT 7159
Qy 2161 GCAGAGTCGAGACCTCTTAAAGGCTTTTAACTTGAAGAAACATTTGCCAGTGCATG 2220
Db 7160 GCAGAGTCGAGACCTCTTAAAGGCTTTTAACTTGAAGAAACATTTGCCAGTGCATG 7219
Qy 2221 ATACCCAAAGCTGAGACCGCTCGCGGACCTGATGATGAAGCAATGCGATGAAACAGAA 2280
Db 7220 ATACCCAAAGCTGAGACCGCTCGCGGACCTGATGATGAAGCAATGCGATGAAACAGAA 7279

2281 TTGGAATTACGACGAGTGTAGTGAAGCCGTAGAAATCCAGATACGAGATCATACTGCGAG 2340
7280 TTGGAATTACGACGAGTGTAGTGAAGCCGTAGAAATCCAGATACGAGATCATACTGCGAG 7339
2341 GCTTATCATCATCGTCTCTGTCCACGTTAGCCGAAAGCGTTAAGAACTTCAAGAGCATTA 2400
7340 GCTTATCATCATCGTCTCTGTCCACGTTAGCCGAAAGCGTTAAGAACTTCAAGAGCATTA 7399
2401 GAGGAGGCCAATCATCCCTTACGCGTGAACCTTAATAGGTGACGTATGATGACCACTT 2460
7400 GAGGAGGCCAATCATCCCTTACGCGTGAACCTTAATAGGTGACGTATGATGACCACTT 7459
2461 ACCCAGCGGCAAGATGTTTCCATACCTTACGCTGAACCTTCCACGAGTTTACCTTACAA 2520
7460 ACCCAGCGGCAAGATGTTTCCATACCTTACGCTGAACCTTCCACGAGTTTACCTTACAA 7519
2521 TCCGATGGCTTACCGAGATCCAAACCTCTTACGCGCGCTGAGGCGCTTTCGCGCC 2580
7520 TCCGATGGCTTACCGAGATCCAAACCTCTTACGCGCGCTGAGGCGCTTTCGCGCC 7579
2581 GCTGCTGCTCAATCGAAGATCTTACGAGGTGATGATGATGATGATGATGATGATGATG 2640
7580 GCTGCTGCTCAATCGAAGATCTTACGAGGTGATGATGATGATGATGATGATGATGATG 7639
2641 ATCACTTATCCGCGCGCGAGTCCACCGGCAAGAAAGAGAGTGTCTTAAAGCAAA 2700
7640 ATCACTTATCCGCGCGCGAGTCCACCGGCAAGAAAGAGAGTGTCTTAAAGCAAA 7699
2701 ACTTACTAGCTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
7700 ACTTACTAGCTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7759
2761 AGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
7760 AGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7819
2821 GAAAGGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2880
7820 GAAAGGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7879
2881 GCTTGAAGAAATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940
7880 GCTTGAAGAAATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7939
2941 GTACACATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3000
7940 GTACACATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7999
3001 CAGGACAAACCAACCGGCGCTTCTACAACTGGGACCAACCGGCGGATGATGATGATGATG 3060
8000 CAGGACAAACCAACCGGCGCTTCTACAACTGGGACCAACCGGCGGATGATGATGATGATG 8059
3061 GAGATTTTACCGTACCGAGAGAGTGGCGGAGAAAGCGACAGCGGAGAACCGATCTTGA 3120
8060 GAGATTTTACCGTACCGAGAGAGTGGCGGAGAAAGCGACAGCGGAGAACCGATCTTGA 8119
3121 CAACGAGGAGAGATGTTGAGTATGTTTGAAGAGTGAATGAGGAGCGGATGATGATGATG 3180
8120 CAACGAGGAGAGATGTTGAGTATGTTTGAAGAGTGAATGAGGAGCGGATGATGATGATG 8179
3181 GCTTCAAGTGTCACTTGGAGCAAGAAAGGAGTGAACATTTAGGATGATGATGATGATGATG 3240
8180 GCTTCAAGTGTCACTTGGAGCAAGAAAGGAGTGAACATTTAGGATGATGATGATGATGATG 8239
3241 TGAACCGTGTCTACTAGTACGCGTATGCGTATGCGTATGCGTATGCGTATGCGTATGCGA 3300
8240 TGAACCGTGTCTACTAGTACGCGTATGCGTATGCGTATGCGTATGCGTATGCGTATGCGA 8299
3301 CAACACACCGTGTGTCTATCTACGCGCAAGAAAGAACTGAGAGTGTGTGAGAGAGAA 3360
8300 CAACACACCGTGTGTCTATCTACGCGCAAGAAAGAACTGAGAGTGTGTGAGAGAGAA 8359

3361 CGTGCAATCCAAATTACGACACGCTGTGAGAGAGCTTTGAATGTCTACACGCGG 3420
8360 CGTGCAATCCAAATTACGACACGCTGTGAGAGAGCTTTGAATGTCTACACGCGG 8419
3421 GCCCAAGAGAGATTTACCGATGACTTCACTGACCACTGCTTCTGAGGCTTCTGCC 3480
8420 GCCCAAGAGAGATTTACCGATGACTTCACTGACCACTGCTTCTGAGGCTTCTGCC 8479
3481 GATTGAGAGACATCAAGCGGCTGTTGAGGCGCAATTTGAATGAGAGAGTGGAGAGA 3540
8480 GATTGAGAGACATCAAGCGGCTGTTGAGGCGCAATTTGAATGAGAGAGTGGAGAGA 8539
3541 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3600
8540 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8599
3601 CACTGCGGATGTCACCAAAATTCGTTACATGTCCTTTCGACCAACGATGATGATGATG 3660
8600 CACTGCGGATGTCACCAAAATTCGTTACATGTCCTTTCGACCAACGATGATGATGATG 8659
3661 AGACGATGAGAGAAATGAGTATGAGACATCTGACCGTGGCGCTTGGCGACAA 3720
8660 AGACGATGAGAGAAATGAGTATGAGACATCTGACCGTGGCGCTTGGCGACAA 8719
3721 AGGTAATCTCTGTTAGCTCAATGTCCTCAAGTGAAGTGTAACTGATGATGATGATG 3780
8720 AGGTAATCTCTGTTAGCTCAATGTCCTCAAGTGAAGTGTAACTGATGATGATGATG 8779
3781 CGAGCATCTGAGAAATTCATGACCGGTGAGAGAAAGATCAGAGAGAGTGTGCTG 3840
8780 CGAGCATCTGAGAAATTCATGACCGGTGAGAGAAAGATCAGAGAGAGTGTGCTG 8839
3841 AGAGAGATCTTGTCCACCGCTCAGTGAAGAGTGTAAAGTGCAGTGAATGATGATG 3900
8840 AGAGAGATCTTGTCCACCGCTCAGTGAAGAGTGTAAAGTGCAGTGAATGATGATG 8899
3901 CTTGAAGAGAGAGTGTGCGGAGTACATGACAGTGAAGAGTGTAAAGTGCAGTGAATG 3960
8900 CTTGAAGAGAGAGTGTGCGGAGTACATGACAGTGAAGAGTGTAAAGTGCAGTGAATG 8959
3961 GTCTTATGAGAGAGAGTGTGAGAGAGAGTGTAAACCACTTGTGCGAGAGAGT 4020
8960 GTCTTATGAGAGAGAGTGTGAGAGAGAGTGTAAACCACTTGTGCGAGAGAGT 9019
4021 CACCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4080
9020 CACCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9079
4081 GAAAGGCTGCACTTAAAGCAAAAGTGTGATGCTTCAAGAGAGAGAGAGAGAGAGAGT 4140
9080 GAAAGGCTGCACTTAAAGCAAAAGTGTGATGCTTCAAGAGAGAGAGAGAGAGAGAGT 9139
4141 CTTCAACTGCGCGGATCTTATTTAGGACACAGACCACTCAAGTGAAGTGAATTTG 4200
9140 CTTCAACTGCGCGGATCTTATTTAGGACACAGACCACTCAAGTGAAGTGAATTTG 9199
4201 TCCATTCGCTTGAACACGAGAGTGTGCGGCTTCCGTTAGTCAACGCTTCAAGTCA 4260
9200 TCCATTCGCTTGAACACGAGAGTGTGCGGCTTCCGTTAGTCAACGCTTCAAGTCA 9259
4261 GAAAGTGTTCAAAGGATCACTTCACTGATGATGATGATGATGATGATGATGATGATG 4320
9260 GAAAGTGTTCAAAGGATCACTTCACTGATGATGATGATGATGATGATGATGATGATG 9319
4321 GAGAAATTTGGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380
9320 GAGAAATTTGGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9379
4381 GAAATTTTCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4440
9380 GAAATTTTCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9439
4441 AGTGGGCGCAGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500

```
Db 9440 AGTGTGGGGCCAGAGAGTGGGACCAAGGAGGACCAATGAGATGGCGATGAGATCATCAT 9499
Qy 4501 CCATCTATATATATGCGCATCCAGTCTACACTGTATGTTGCTGTGTGTGCTCTTGC 4560
Db 9500 CCACATATATATCATGCGGATCCAGTCTACACTGTATGTTGCTGTGTGTGCTCTTGC 9559
Qy 4561 TATCTGTGTGTGCGATGATGATGAGAGCTTGTGATGCGCAAGAACAGAGAGAGAGCTGCT 4620
Db 9560 TATCTGTGTGTGCGATGATGATGAGAGCTTGTGATGCGCAAGAACAGAGAGAGAGCTGCT 9619
Qy 4621 GAGCGCATACGCGCTTGACCCGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 4680
Db 9620 GAGCGCATACGCGCTTGACCCGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 9679
Qy 4681 CATTCGGCGCAACCAAGCTGAAACATTTGGAGAACTTTGAACCATCTGTGTGTTAAACA 4740
Db 9680 CATTCGGCGCAACCAAGCTGAAACATTTGGAGAACTTTGAACCATCTGTGTGTTAAACA 9739
Qy 4741 CCAACCGTTTCTGAGGAGAGTGTGATTCCTGAGGAGCGTGTGTTATCTGTTCCG 4800
Db 9740 CCAACCGTTTCTGAGGAGAGTGTGATTCCTGAGGAGCGTGTGTTATCTGTTCCG 9799
Qy 4801 CTGCTTTTCATGCTGCAATGCTTTTATTTGTTGCAAGGCGTCTGCTGAGGAAAGTAGA 4860
Db 9800 CTGCTTTTCATGCTGCAATGCTTTTATTTGTTGCAAGGCGTCTGCTGAGGAAAGTAGA 9859
Qy 4861 CGCCTTTGAAACATGCAACATCTGTGCAATGTTCCGGGAGTCCCGATTAAGGGCTGTGT 4920
Db 9860 CGCCTTTGAAACATGCAACATCTGTGCAATGTTCCGGGAGTCCCGATTAAGGGCTGTGT 9919
Qy 4921 CGAAGCGGCGAGTTACGCGGCACTTAACCTGAGAGATCGAGTCCGTCATCGGAATTAC 4980
Db 9920 CGAAGCGGCGAGTTACGCGGCACTTAACCTGAGAGATCGAGTCCGTCATCGGAATTAC 9979
Qy 4981 ACCTTCAACTAACAAGAGATGCTGACCTGCAAACTTCCACAGTATCTCTTACACA 5040
Db 9980 ACCTTCAACTAACAAGAGATGCTGACCTGCAAACTTCCACAGTATCTCTTACACA 10039
Qy 5041 AGTTAAATGCTGCGGCTCTCGAGTGAAGGCACTCTCAAGGCGGATTAACATGCGG 5100
Db 10040 AGTTAAATGCTGCGGCTCTCGAGTGAAGGCACTCTCAAGGCGGATTAACATGCGG 10099
Qy 5101 CGTTTTGGGCGGTGATACCTTTCATGTTGGGAGGCGCAATGCTTCTGTGACAGTGA 5160
Db 10100 CGTTTTGGGCGGTGATACCTTTCATGTTGGGAGGCGCAATGCTTCTGTGACAGTGA 10159
Qy 5161 GAACACAACTGAGTGAAGGCGTGAAGTTCGCTTCAGACTGCACTATAGATCAAGC 5220
Db 10160 GAACACAACTGAGTGAAGGCGTGAAGTTCGCTTCAGACTGCACTATAGATCAAGC 10219
Qy 5221 AGTGCACATAAAGTTCAACAGCTGCTTGAAGTGGCGCTGCGTATAGATACGGCA 5280
Db 10220 AGTGCACATAAAGTTCAACAGCTGCTTGAAGTGGCGCTGCGTATAGATACGGCA 10279
Qy 5281 CACCAACGCGGACCTGATAGCTTGTCAATGGGCTACGCGAGTTCTCCACGGGACCT 5340
Db 10280 CACCAACGCGGACCTGATAGCTTGTCAATGGGCTACGCGAGTTCTCCACGGGACCT 10339
Qy 5341 GAAGGTATAGCAGGCGGATATGAGCGCTTTTTCACCTTTGACATAGAGTGTGAT 5400
Db 10340 GAAGGTATAGCAGGCGGATATGAGCGCTTTTTCACCTTTGACATAGAGTGTGAT 10399
Qy 5401 CAGAAAGGGGCTTTTTCACATAGCTTCTCGAGTATGAGGATAGAAACAGAGAGC 5460
Db 10400 CAGAAAGGGGCTTTTTCACATAGCTTCTCGAGTATGAGGATAGAAACAGAGAGC 10459
Qy 5461 GTTGGGAGATTTAAGGATCTGCTTGAAGTATGAGTATGAGGATAGGAGGAGGAGAT 5520
Db 10460 GTTGGGAGATTTAAGGATCTGCTTGAAGTATGAGTATGAGGATAGGAGGAGGAGAT 10519
Qy 5521 ACGGCTCTGAAGCTTCTGTCAAGAACATCCAGTCCCTTACACCAAGAGATATCAG 5580
Db 10520 ACGGCTCTGAAGCTTCTGTCAAGAACATCCAGTCCCTTACACCAAGAGATATCAG 10579
Qy 5581 GTATGAATGTGAAAGAACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5640
Db 10580 GTATGAATGTGAAAGAACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10639
Qy 5641 TAAATTTGAAGTGAAGCTTCTGAGAGGCTTAACTGTCTTACGAGGAGACATCCATCTC 5700
Db 10640 TAAATTTGAAGTGAAGCTTCTGAGAGGCTTAACTGTCTTACGAGGAGACATCCATCTC 10699
Qy 5701 GATTGACATCTCTGATGACCTTTTGTGATCATCAAGATCAACAAATTTAGAGT 5760
Db 10700 GATTGACATCTCTGATGACCTTTTGTGATCATCAAGATCAACAAATTTAGAGT 10759
Qy 5761 TAGCTGACAGTATGACAGTCTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5820
Db 10760 TAGCTGACAGTATGACAGTCTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10819
Qy 5821 GTACAAAGCTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5880
Db 10820 GTACAAAGCTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10879
Qy 5881 GAAGAAAGGACCAACATGATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5940
Db 10880 GAAGAAAGGACCAACATGATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10939
Qy 5941 GAGCGCAAGGCAATTTTATGATTTGCTATGCGGAGAGAGAGAGAGAGAGAGAGAGAG 6000
Db 10940 GAGCGCAAGGCAATTTTATGATTTGCTATGCGGAGAGAGAGAGAGAGAGAGAGAGAG 10999
Qy 6001 ATGTAAACACCGGCGGACCAACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6060
Db 11000 ATGTAAACACCGGCGGACCAACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11059
Qy 6061 GCGGCGAGTTTCCAAACATCTTGAAGCTGCTGCTTGAAGCTGCTGCTGCTGCTGCTGCT 6120
Db 11060 GCGGCGAGTTTCCAAACATCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11119
Qy 6121 CCTCATTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6180
Db 11120 CCTCATTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11179
Qy 6181 ATGACTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6240
Db 11180 ATGACTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11239
Qy 6241 TGCATTAATGCGAGCGGCGCTTGAACATTAACCTGATGATTTCCGAGAGAGAGAGAG 6300
Db 11240 TGCATTAATGCGAGCGGCGCTTGAACATTAACCTGATGATTTCCGAGAGAGAGAGAG 11299
Qy 6301 CATTAATGCTGTGAGTGTCACTTAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6360
Db 11300 CATTAATGCTGTGAGTGTCACTTAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11359
Qy 6361 CACTTTATGAGACTCACTATGAGTGTCTATATATACACTACATATTTACTTTAAAC 6420
Db 11360 CACTTTATGAGACTCACTATGAGTGTCTATATATACACTACATATTTACTTTAAAC 11419
Qy 6421 ACTATACACTTATTAATTTCTTTTATTTTCTTTTGTTTTATTTTGTTTTAA 6480
Db 11420 ACTATACACTTATTAATTTCTTTTATTTTCTTTTGTTTTATTTTGTTTTAA 11479
Qy 6481 ATTTTC 6485
Db 11480 ATTTTC 11484
```

RESULT 2
US-10-023-649-7
; Sequence 7, Application US/10023649
; Publication No. US20030143201A1
; GENERAL INFORMATION:
; APPLICANT: The Minister of National Defence, Government of Canada

APPLICANT: Nagata, Leslie P
APPLICANT: Wong, Jonathan P
TITLE OF INVENTION: No. US20030143201A1el DNA-Based Vaccine Against the Encephalitis
FILE REFERENCE: NEL-001
CURRENT APPLICATION NUMBER: US/10/023,649
CURRENT FILING DATE: 2002-12-27
PRIOR APPLICATION NUMBER: 60/256,948
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 4395
TYPE: DNA
ORGANISM: western equine encephalomyelitis virus - STRAIN 71v-1658
FEATURE:
NAME/KEY: CMV promoter
LOCATION: (1)..(1260)
OTHER INFORMATION: PvaX vector sequence: 1-196; CMV promoter: 1-115; CMV putative tr
OTHER INFORMATION: anscritpional start site: 125; T7 promoter: 48-167; PVAx multicld
OTHER INFORMATION: ning region: 166-196; polypotein (C-E3-E2-6K-E1): 214-4065; pCDM
OTHER INFORMATION: -HX45 nontranslated region: 4066-4348; pCDM-HX45 vector sequence:
US-10-023-649-7

Query Match 63.8%; Score 4137.2; DB 12; Length 4395;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4139; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2324 CGAATCATCTGCGAGGCTGATCATCATGCTCTTGTCCACGTTAGCCGAAAGCGTTAA 2383
DB 207 CGAGCTATATCTGCGAGGCTGATCATCATGCTCTTGTCCACGTTAGCCGAAAGCGTTAA 266
QY 2384 GAATCTCAAGACATTAAGAGGAGGCCAATCACTCTAGCGGCTGACCTAAATAGGTGAC 2443
DB 267 GAATCTCAAGACATTAAGAGGAGGCCAATCACTCTAGCGGCTGACCTAAATAGGTGAC 326
QY 2444 GTATAGACAGCACTTACCCACCGGCAAGTGTTCATACCTCAGCTGAACTTTTCA 2503
DB 327 GTATAGACAGCACTTACCCACCGGCAAGTGTTCATACCTCAGCTGAACTTTTCA 386
QY 2504 CCAATTTACCTTACCAATTCGATGCTTACCGAATCCAACTCTAGGCGCGCTG 2563
DB 387 CCAATTTACCTTACCAATTCGATGCTTACCGAATCCAACTCTAGGCGCGCTG 446
QY 2564 AGGCGCTTGGCCCGGCTGCTGCTCAATTCGAATCTTGAAGGTGATGTCAC 2623
DB 447 AGGCGCTTGGCCCGGCTGCTGCTCAATTCGAATCTTGAAGGTGATGTCAC 506
QY 2624 TTGACTTTCAACACGATCACTAATCCGCGCAGGTCCACCGCAAGAAAGAG 2683
DB 507 TTGACTTTCAACACGATCACTAATCCGCGCAGGTCCACCGCAAGAAAGAG 566
QY 2684 AGTCTCTTAAAGCAAACTTAATCTGATGCTTAAAGAAAGAGCAAGCAAGAGAG 2743
DB 567 AGTCTCTTAAAGCAAACTTAATCTGATGCTTAAAGAAAGAGCAAGCAAGAGAG 626
QY 2744 AAAGCAAGCTTAAAGCAAGGAAAGCAAGTATGCTAATGAAGTTGATGCTGGAACAG 2803
DB 627 AAAGCAAGCTTAAAGCAAGGAAAGCAAGTATGCTAATGAAGTTGATGCTGGAACAG 686
QY 2804 ACATTTTCGATCATGCTGAAAGGCAAGTGAATGATGCTGCTGCTGCTGGAAGAG 2863
DB 687 ACATTTTCGATCATGCTGAAAGGCAAGTGAATGATGCTGCTGCTGCTGGAAGAG 746
QY 2864 CTGATGAAACCACTTCACTGCTGAAAGAAATTTGATTAAGCAATTAGCGCGGTGAA 2923
DB 747 CTGATGAAACCACTTCACTGCTGAAAGAAATTTGATTAAGCAATTAGCGCGGTGAA 806
QY 2924 TTGAAGAGGCTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2983
DB 807 TTGAAGAGGCTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 866
QY 2984 GACACGCTGACGTAACACGAGCAAAACACCGGCTTCTTCAACTGACCAACGAGCGCA 3043

DB 867 GACACGCTGACGTAACACGAGCAAAACACCGGCTTCTTCAACTGACCAACGAGCGCA 926
QY 3044 GTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3103
DB 927 GTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 986
QY 3104 GGAAGACGATCTGACCAAGAGGAGAGTGTGCTAATTTGTTCTGAGAGGTCAAT 3163
DB 987 GGAAGACGATCTGACCAAGAGGAGAGTGTGCTAATTTGTTCTGAGAGGTCAAT 1046
QY 3164 GAGGACGCGCTACGCGCTTTCACTGCTCACTTGAACCAAGAAAGGCTGATGAG 3223
DB 1047 GAGGACGCGCTACGCGCTTTCACTGCTCACTTGAACCAAGAAAGGCTGATGAG 1106
QY 3224 GATACCCCCGAAAGTTCTGAACCGTGTCACTAGTACAGGCTATGCTGCTTCAAT 3283
DB 1107 GATACCCCCGAAAGTTCTGAACCGTGTCACTAGTACAGGCTATGCTGCTTCAAT 1166
QY 3284 GTCAAGTTCCATGCGACCAACCGGCTGCTATTCATGAGCGCGCAAGCAACATC 3343
DB 1167 GTCAAGTTCCATGCGACCAACCGGCTGCTATTCATGAGCGCGCAAGCAACATC 1226
QY 3344 GACGTGCTCGAAGAAAGTTCGACCAATTCACACGCTGCTGAGAACGCTTTG 3403
DB 1227 GACGTGCTCGAAGAAAGTTCGACCAATTCACACGCTGCTGAGAACGCTTTG 1286
QY 3404 AAATGTCATACCGCGCGCCCAACCAAGCACTTACCGATGCTTCACTGACCACTCC 3463
DB 1287 AAATGTCATACCGCGCGCCCAACCAAGCACTTACCGATGCTTCACTGACCACTCC 1346
QY 3464 TACCTGGGGTCTGCGCGATTTGACACCAAGCGCGTTCAGCGCAATTAATTAAT 3523
DB 1347 TACCTGGGGTCTGCGCGATTTGACACCAAGCGCGTTCAGCGCAATTAATTAAT 1406
QY 3524 GAGAAAGTGTGGAAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 3583
DB 1407 GAGAAAGTGTGGAAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1466
QY 3584 GGTCTACATACGAGGACCTGCGGATGTCAACCAATTCGCTTACATGTCTTTGACAC 3643
DB 1467 GGTCTACATACGAGGACCTGCGGATGTCAACCAATTCGCTTACATGTCTTTGACAC 1526
QY 3644 GACCATGACATCAAGGAAGACATGATGAGAAATGATGATGATGATGATGATGATGATGAT 3703
DB 1527 GACCATGACATCAAGGAAGACATGATGAGAAATGATGATGATGATGATGATGATGATGAT 1586
QY 3704 GGTGCTTGGCCCAAAAGGCTACTTCTGTTAGCTCAATGCTCTCCAGGTGACAGTGA 3763
DB 1587 GGTGCTTGGCCCAAAAGGCTACTTCTGTTAGCTCAATGCTCTCCAGGTGACAGTGA 1646
QY 3764 ACCGTACATACGAGCGGACATCTGAGATTCATGACCGGTGAGAGAAAGATCAGG 3823
DB 1647 ACCGTACATACGAGCGGACATCTGAGATTCATGACCGGTGAGAGAAAGATCAGG 1706
QY 3824 AGAAGTTGTGCTGATGAGAGATCTGTTCCACCGGCTCAATGAAAGCTGTAAG 3883
DB 1707 AGAAGTTGTGCTGATGAGAGATCTGTTCCACCGGCTCAATGAAAGCTGTAAG 1766
QY 3884 TGCCAGCTTACGATCACTTGAAGAGACGCTGCGGGTACATTAACCATGACAGGCA 3943
DB 1767 TGCCAGCTTACGATCACTTGAAGAGACGCTGCGGGTACATTAACCATGACAGGCA 1826
QY 3944 GGCCTCACCGCTTAACTCTATCTGAGAGAGCGTCAAGCGAAGTGTACATTAACCA 4003
DB 1827 GGCCTCACCGCTTAACTCTATCTGAGAGAGCGTCAAGCGAAGTGTACATTAACCA 1886
QY 4004 CTTTCTGCAAGACGTCACCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4063
DB 1887 CTTTCTGCAAGACGTCACCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1946
QY 4064 AGCAGCGAAGAGATGAAAGGCTGCACTTAAGCAAAACAGTGCATTTGCTTCAAGAGC 4123

Dh 1947 AGCAGCGAAGCAAGATGAAAGCGCTGCACTAAAGCAAAAGATGTCATTGGCTTCAAGAGC 2006
Qy 4124 GACCAAGCAAAAGGGCTTCACTCGCCGGAATCTTATAGGCAACAGACCACTCGATG 4183
Dh 2007 GACCAAGCAAAAGGGCTTCACTCGCCGGAATCTTATAGGCAACAGACCACTCGATG 2066
Qy 4184 CAAGGTAAATTGCACTTCCATTCGGCTTGAACCGACAGCTGCGCGGTTCGGTTAGCT 4243
Dh 2067 CAAGGTAAATTGCACTTCCATTCGGCTTGAACCGACAGCTGCGCGGTTCGGTTAGCT 2126
Qy 4244 CACACGCTTCACTGCAAGAGGTGTTCAAGGCACTCACTCCACTGACTGCAATCGA 4303
Dh 2127 CACACGCTTCACTGCAAGAGGTGTTCAAGGCACTCACTCCACTGACTGCAATCGA 2186
Qy 4304 CCAACATGCTGCAAGAGGAAATGGGGCTGCGACAGACCGAAGAGCAATGGAAT 4363
Dh 2187 CCAACATGCTGCAAGAGGAAATGGGGCTGCGACAGACCGAAGAGCAATGGAAT 2246
Qy 4364 ACAGGCTCTACATCAAGAAATTTTCTGTGGGCGAAGAGGCTGAGTACGTATGGGT 4423
Dh 2247 ACAGGCTCTACATCAAGAAATTTTCTGTGGGCGAAGAGGCTGAGTACGTATGGGT 2306
Qy 4424 AACCATGAACAGTCAAGTCTGGGGCCAGAGAGTGGCAACAGGCAACCAATGATGG 4483
Dh 2307 AACCATGAACAGTCAAGTCTGGGGCCAGAGAGTGGCAACAGGCAACCAATGATGG 2366
Qy 4484 CCGATGAGATCAATCACTCACTATATCATGGGATCAGTCTACATGTCATGTGCTG 4543
Dh 2367 CCGATGAGATCAATCACTCACTATATCATGGGATCAGTCTACATGTCATGTGCTG 2426
Qy 4544 TGTGCTGCTCTTGTGCTATCTGTGAGGCACTGCACTATCAGAGCTTGATGCGCAA 4603
Dh 2427 TGTGCTGCTCTTGTGCTATCTGTGAGGCACTGCACTATCAGAGCTTGATGCGCAA 2486
Qy 4604 GCAAGAAAGATCTGCTGAGCGCCATACGCGCTTGCAACGAAACGAAACGTAACCAAGCA 4663
Dh 2487 GCAAGAAAGATCTGCTGAGCGCCATACGCGCTTGCAACGAAACGTAACCAAGCA 2546
Qy 4664 TTAGCGGTTTGTGCTGCTATCTGCGCAACCAAGCTGAAACATTTGAGAACTTTGAAAC 4723
Dh 2547 TTAGCGGTTTGTGCTGCTATCTGCGCAACCAAGCTGAAACATTTGAGAACTTTGAAAC 2606
Qy 4724 CATCTGTGTTTAAACAACAACGCTTCTCTGCGCAGAGTGTGCTTCTCTGCGCAGC 4783
Dh 2607 CATCTGTGTTTAAACAACAACGCTTCTCTGCGCAGAGTGTGCTTCTCTGCGCAGC 2666
Qy 4784 CTGTGTTTCTGCTGCTGCTTCTTCACTGCTGCACTGCTTTTATTTGTTGCGCGCTC 4843
Dh 2667 CTGTGTTTCTGCTGCTGCTTCTTCACTGCTGCACTGCTTTTATTTGTTGCGCGCTC 2726
Qy 4844 TGCCTGGGGAAGGTAAAGCGCTTGAACATGCGACCACTGCGCAAAATGTTCCGGGATC 4903
Dh 2727 TGCCTGGGGAAGGTAAAGCGCTTGAACATGCGACCACTGCGCAAAATGTTCCGGGATC 2786
Qy 4904 CCGATTAAGGCTGTGCTGAACGCGCAGTTACCGGCACTTAACTGAGATCAAGCTC 4963
Dh 2787 CCGATTAAGGCTGTGCTGAACGCGCAGTTACCGGCACTTAACTGAGATCAAGCTC 2846
Qy 4964 GTCTCATCGGAATTAACACTTCACTTAACAAGAGTACGTGACCTGCAAAATTCACACA 5023
Dh 2847 GTCTCATCGGAATTAACACTTCACTTAACAAGAGTACGTGACCTGCAAAATTCACACA 2906
Qy 5024 GTCACTTCTTCAACAAGTTAAATGCTGCGGGTCCCTCGAGTCAAGGCACTCTCAAG 5083
Dh 2907 GTCACTTCTTCAACAAGTTAAATGCTGCGGGTCCCTCGAGTCAAGGCACTCTCAAG 2966
Qy 5084 GCGGATTAACATGCGCGGTTTTTGGCGGTGTGTACCTTTCACTGTTGGGAGCGCAAA 5143
Dh 2967 GCGGATTAACATGCGCGGTTTTTGGCGGTGTGTACCTTTCACTGTTGGGAGCGCAAA 3026
Qy 5144 TGCTTCTGTGACAGTGAACAACAACAATGAGTGAAGGCGTACGTGAGTTGCTTCAAGC 5203
Dh 3027 TGCTTCTGTGACAGTGAACAACAACAATGAGTGAAGGCGTACGTGAGTTGCTTCAAGC 3086

Qy 5204 TGCATATAGATCAGCAGTCCGACTTAAAGTTCAACACAGCTGCTGAAAGTCGGCTG 5263
Dh 3087 TGCATATAGATCAGCAGTCCGACTTAAAGTTCAACACAGCTGCTGAAAGTCGGCTG 3146
Qy 5264 CGTATAGTATACGGCAACACACCGCGCACTGGAATCGTTTGTCAATGGCGTCAAGCA 5323
Dh 3147 CGTATAGTATACGGCAACACACCGCGCACTGGAATCGTTTGTCAATGGCGTCAAGCA 3306
Qy 5324 GGTTCCTCAAGGGAACCTGAAGGTCAATGAGGCGGCAATACACCGCTTTTACCCCTT 5383
Dh 3207 GGTTCCTCAAGGGAACCTGAAGGTCAATGAGGCGGCAATACACCGCTTTTACCCCTT 3266
Qy 5384 GACCATTAAGTCTGCTATCAGAAAGGCGCTGTTTAACTACACTGACTTCCCTGAGTATGA 5443
Dh 3267 GACCATTAAGTCTGCTATCAGAAAGGCGCTGTTTAACTACACTGACTTCCCTGAGTATGA 3326
Qy 5444 GCTATGAACACAGAGCGTTCCGCAATATTCAAGCATCTCTGCTTGAATGCTAACAGATA 5503
Dh 3327 GCTATGAACACAGAGCGTTCCGCAATATTCAAGCATCTCTGCTTGAATGCTAACAGATA 3386
Qy 5504 GTAGCCCGCACTGACATACCGCTGCTGAAGCTTCTGTCAAGAACATCCACGTCCTTAC 5563
Dh 3387 GTAGCCCGCACTGACATACCGCTGCTGAAGCTTCTGTCAAGAACATCCACGTCCTTAC 3446
Qy 5564 ACCCAAGCAGTATCAGGCTATGAATGTGGAAGAACACTCAGAGACGACCCCTGCAAGAA 5623
Dh 3447 ACCCAAGCAGTATCAGGCTATGAATGTGGAAGAACACTCAGAGACGACCCCTGCAAGAA 3506
Qy 5624 ACAGACCAATTTGATGATTAATTAATGAATGAGAGCTCTGCGACGCTTAACTGTCTTAC 5683
Dh 3507 ACAGACCAATTTGATGATTAATTAATGAATGAGAGCTCTGCGACGCTTAACTGTCTTAC 3566
Qy 5684 GGGCAATCCCTATCTCGATTTGAATCCCTGATGAGAGCTTTGAGATATACAGAAATCA 5743
Dh 3567 GGGCAATCCCTATCTCGATTTGAATCCCTGATGAGAGCTTTGAGATATACAGAAATCA 3626
Qy 5744 CCAACAATTTTAAGATTAGCTGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5803
Dh 3627 CCAACAATTTTAAGATTAGCTGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3686
Qy 5804 GGTTCCTTAACTTAACTGAATCAAAAGCTGACAGGAGGAGCAATTTGCTACTCCAC 5863
Dh 3687 GGTTCCTTAACTTAACTGAATCAAAAGCTGACAGGAGGAGCAATTTGCTACTCCAC 3746
Qy 5864 TCCACGACAGCTGTTTGAAGGAGCAACACATGTAATCCCTGAGGAGGATACA 5923
Dh 3747 TCCACGACAGCTGTTTGAAGGAGCAACACATGTAATCCCTGAGGAGGATACA 3806
Qy 5924 CTACATTTTGAACATGAGCCCAACAAGCAAAATTTTATGTTTCTGCTATCGGCAAGAG 5983
Dh 3807 CTACATTTTGAACATGAGCCCAACAAGCAAAATTTTATGTTTCTGCTATCGGCAAGAG 3866
Qy 5984 TCCACTGCAATGCTGAATGTAACCAACGCGCGACCACTAAATTTGAGAACCAATPAA 6043
Dh 3867 TCCACTGCAATGCTGAATGTAACCAACGCGCGACCACTAAATTTGAGAACCAATPAA 3926
Qy 6044 GTGACCTAAGAAATTCAGGCGGCAAGTTTCAAAACATCTTGAACCTGCTTGAACCTG 6103
Dh 3927 GTGACCTAAGAAATTCAGGCGGCAAGTTTCAAAACATCTTGAACCTGCTTGAACCTG 3986
Qy 6104 TTTGGGGGAGCATCATCTCTCATTTGTTGATGAGCTTATAGTGTGCTCAGCTCATG 6163
Dh 3987 TTTGGGGGAGCATCATCTCTCATTTGTTGATGAGCTTATAGTGTGCTCAGCTCATG 4046
Qy 6164 CTTATATAACACAGTATGATGATGAGCGCGAACAATGACATGACGATTAATTCGATGTA 6223
Dh 4047 CTTATATAACACAGTATGATGATGAGCGCGAACAATGACATGACGATTAATTCGATGTA 4106
Qy 6224 CTTCCGAGGAAGGTGTGATTAATGCGACGCGCTTGAACATTAATTCGATGATTA 6283
Dh 4107 CTTCCGAGGAAGGTGTGATTAATGCGACGCGCTTGAACATTAATTCGATGATTA 4166

QY 6294 TCCGAGGAGCAGCTGATTAATGCTGTGACAGTGCACATTAATGTAATACACTACA 6343
DB 4167 TCCGAGGAGCAGCTGATTAATGCTGTGACAGTGCACATTAATGTAATACACTACA 4226
QY 6344 TATTAAACAACCTAATATCACTTTTATGAGACTCACTATGAGTCTTAATATACACTACAC 6403
DB 4227 TATTAAACAACCTAATATCACTTTTATGAGACTCACTATGAGTCTTAATATACACTACAC 4286
QY 6404 ATATTTCCTTAAAAAAGCTATACACCTTTTAAATCTTTTATATTTTCTTTGTT 6463
DB 4287 ATATTTCCTTAAAAAAGCTATACACCTTTTAAATCTTTTATATTTTCTTTGTT 4346
QY 6464 TT 6465
DB 4347 TT 4348

RESULT 3

US-10-023-649-5
; Sequence 5, Application US/10023649
; Publication No. US20030143201A1
; GENERAL INFORMATION:
; APPLICANT: The Minister of National Defence, Government of Canada
; APPLICANT: Negata, Leslie P
; APPLICANT: Wong, Jonathan P
; TITLE OF INVENTION: No. US20030143201A1e1 DNA-Based Vaccine Against the Encephalitis
; FILE REFERENCE: NEL-001
; CURRENT APPLICATION NUMBER: US/10/023,649
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: 60/256,948
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 4150
; TYPE: DNA
; ORGANISM: Western equine encephalomyelitis virus - strain 71V-1658
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)..(3869)
; OTHER INFORMATION: vector sequence 1-9; 5' SacI primer 9-20; 3' end of NS4 gene 16-1
; OTHER INFORMATION: 14; intergenic region 115-158; polypeptide (C-E3-E2-6K-E1) 159-38
; OTHER INFORMATION: 56; pCDM-XH7 nontranslated region 3857-4150
US-10-023-649-5

Query Match 63.7%; Score 4129.8; DB 12; Length 4150;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4134; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2325 GAGATCATCTGAGGAGGCTGATCATCATCTGTCACAGTTAGCCGAAAGCGTTAAG 2384
DB 10 GAGCTCAATCTGAGGAGGCTGATCATCATCTGTCACAGTTAGCCGAAAGCGTTAAG 69
QY 2385 AACTTCAAGAGCATTAAGAGGAGGCCAATCACTCTCTAAGGCTGAAGCTTAATAGGTGAG 2444
DB 70 AACTTCAAGAGCATTAAGAGGAGGCCAATCACTCTCTAAGGCTGAAGCTTAATAGGTGAG 129
QY 2445 TAGTGAACACGACCTTACCCACCGGAGAAATTTTTCATACCTTAAGCTGAAGCTTTCCAC 2504
DB 130 TAGTGAACACGACCTTACCCACCGGAGAAATTTTTCATACCTTAAGCTGAAGCTTTCCAC 189
QY 2505 CAGTTTACCTTACCAATCCGATGCTTACCGAGATCCAAACCTCTCTAGGCGCGCTGGA 2564
DB 190 CAGTTTACCTTACCAATCCGATGCTTACCGAGATCCAAACCTCTCTAGGCGCGCTGGA 249
QY 2565 GAGCGTTTGGCGCGCGCTGCTGCTCAATTCGAAGATCTTGAAGAGTCAATGTAAGTCACT 2624
DB 250 GAGCGTTTGGCGCGCGCTGCTGCTCAATTCGAAGATCTTGAAGAGTCAATGTAAGTCACT 309
QY 2625 TGACTTTCAACACGATCACTTAATCCGCGCGCGAGTTCACCGCCAAAGAAAGAAAGA 2684
DB 310 TGACTTTCAACACGATCACTTAATCCGCGCGCGAGTTCACCGCCAAAGAAAGAAAGA 369

QY 2685 GTGCTCTTAAGCAAAACCTTACTGACCTTAAAAAGAAAGAGCAAGCAAGAGAGA 2744
DB 370 GTGCTCTTAAGCAAAACCTTACTGACCTTAAAAAGAAAGAGCAAGCAAGAGAGA 429
QY 2745 AACGCAAGCTTAAACGAGGAAACGACAGATGTATGAAAGTTGAGATCGGACAAAGA 2804
DB 430 AACGCAAGCTTAAACGAGGAAACGACAGATGTATGAAAGTTGAGATCGGACAAAGA 489
QY 2805 CATTTCCGATGATGCTGAACGGCCGAAGTAATGATAAGCTGCTTCCGAGAGAGAGC 2864
DB 490 CATTTCCGATGATGCTGAACGGCCGAAGTAATGATAAGCTGCTTCCGAGAGAGAGC 549
QY 2865 TGATGAACCACTCCACCTTGAAGGAAATTTGATTAATGACCAATTTAGCCGCCGTGAAT 2924
DB 550 TGATGAACCACTCCACCTTGAAGGAAATTTGATTAATGACCAATTTAGCCGCCGTGAAT 609
QY 2925 TGAAGAGGCTTACGATGTAAGCTTGAAGTACGAGGAGTTCCTCCAGAACATGAATCAG 2984
DB 610 TGAAGAGGCTTACGATGTAAGCTTGAAGTACGAGGAGTTCCTCCAGAACATGAATCAG 669
QY 2985 ACAGCTGACATTAACCAAGGAGCAAAACCAAGGAGTTCCTTACCACTGAGCAAGCGGAG 3044
DB 670 ACAGCTGACATTAACCAAGGAGCAAAACCAAGGAGTTCCTTACCACTGAGCAAGCGGAG 729
QY 3045 TCCAGTATGAGAAATGAGAGATTTTACCGTACCGAGAGAGTGGCGGAAAGCGACAGCG 3104
DB 730 TCCAGTATGAGAAATGAGAGATTTTACCGTACCGAGAGAGTGGCGGAAAGCGACAGCG 789
QY 3105 GAAGACCGATCTTGACACCAAGAGCAGAGTGTGCTATTGTTCTAGAGGTGCAATG 3164
DB 790 GAAGACCGATCTTGACACCAAGAGCAGAGTGTGCTATTGTTCTAGAGGTGCAATG 849
QY 3165 AGGGAAGCGCTTACGAGCGCTTTCAGTGTGCTGCTGGAACCAAGAAAGGCTGACATTAAG 3224
DB 850 AGGGAAGCGCTTACGAGCGCTTTCAGTGTGCTGCTGGAACCAAGAAAGGCTGACATTAAG 909
QY 3225 ATACCCCGAGAGTCTTGAACCGTGTGCTAGTATTAACGAGCTATGCTTTGCAATG 3284
DB 910 ATACCCCGAGAGTCTTGAACCGTGTGCTAGTATTAACGAGCTATGCTTTGCAATG 969
QY 3285 TCAGTGTCCATGCGACAAACCAAGCCTGCTATTTCACTGACCGCAGAAAGCAACATCTG 3344
DB 970 TCAGTGTCCATGCGACAAACCAAGCCTGCTATTTCACTGACCGCAGAAAGCAACATCTG 1029
QY 3345 ACGTCTGAGAGAAACGTCGACAAATCCAAATTCGACACGCTGTGAGAAAGCTTGA 3404
DB 1030 ACGTCTGAGAGAAACGTCGACAAATCCAAATTCGACACGCTGTGAGAAAGCTTGA 1089
QY 3405 AATGTCATCAAGCGCGGCCCAAGCAAGCATTTACGATGATCTTCACTGACAGTCCCT 3464
DB 1090 AATGTCATCAAGCGCGGCCCAAGCAAGCATTTACGATGATCTTCACTGACAGTCCCT 1149
QY 3465 ACTGGGGTTTGGCCGATTTGACAGACCTCAAGCGCTGTTTCAGCCCAATAAATTTG 3524
DB 1150 ACTGGGGTTTGGCCGATTTGACAGACCTCAAGCGCTGTTTCAGCCCAATAAATTTG 1209
QY 3525 AGAAGTGTGAGAGAAATCTATGATGATGATGATTAAGATTCAGAGTCTCGGACAAATTCG 3584
DB 1210 AGAAGTGTGAGAGAAATCTATGATGATGATGATTAAGATTCAGAGTCTCGGACAAATTCG 1269
QY 3585 GCTACAAATCAGAGGAGCCTGCGGATGTCACCAATTCGTTACATGCTTTTCGACAG 3644
DB 1270 GCTACAAATCAGAGGAGCCTGCGGATGTCACCAATTCGTTACATGCTTTTCGACAG 1329
QY 3645 ACCATGACATCAGAGAGACAGTATGAGAAATAGCTATCAGCAATCTGACCCCTG 3704
DB 1330 ACCATGACATCAGAGAGACAGTATGAGAAATAGCTATCAGCAATCTGACCCCTG 1389
QY 3705 GTGCTTTGGGCAAAAGGATCTTCTGTTAGCTCAATGTCCTTCAGGTGACAGTGA 3764
DB 1390 GTGCTTTGGGCAAAAGGATCTTCTGTTAGCTCAATGTCCTTCAGGTGACAGTGA 1449
QY 3765 CCGTCAGTATCAGAGGCGAGCATCTGAGATTCATGACCGTGGAGAAAGATCAAGA 3824

```
Db 1450 CCGTCAATACAGAGGAGGAGCATCTGAAATTCATCAACGTGAGAAAAAGATCAAGA 1509
Qy 3825 GGAAGTTTGTGCGTAGAGAGAGATCTTTTCCACCCCTTCATGAAAGTGGTAAAGT 3884
Db 1510 GGAAGTTTGTGCGTAGAGAGAGATCTTTTCCACCCCTTCATGAAAGTGGTAAAGT 1569
Qy 3885 GCCACGTTTACATCACTTGAAGGAGAGCTCTGCGGGTACATPAAACATGACAGGCGAG 3944
Db 1570 GCCACGTTTACATCACTTGAAGGAGAGCTCTGCGGGTACATPAAACATGACAGGCGAG 1629
Qy 3945 GCCCAACGCGTAAAGTCTTATCTGAGAGAGAGGCTGAGGCGAAGTGTACATTTAAACAC 4004
Db 1630 GCCCAACGCGTAAAGTCTTATCTGAGAGAGAGGCTGAGGCGAAGTGTACATTTAAACAC 1689
Qy 4005 CTTTGGCAAGAAAGTCACTCAATGTAAAGTGTGCGACTACAGCAAGGTATCGTGA 4064
Db 1690 CTTTGGCAAGAAAGTCACTCAATGTAAAGTGTGCGACTACAGCAAGGTATCGTGA 1749
Qy 4065 GCAACGGAAGAAAGTGAAGGCTGCACTAAAGCAAAAGTGAATTGCCCTAACAAAGAGG 4124
Db 1750 GCAACGGAAGAAAGTGAAGGCTGCACTAAAGCAAAAGTGAATTGCCCTAACAAAGAGG 1809
Qy 4125 ACCAAAGAAATGGGCTTCAACTCGCGAGATCTTATTAAGCAACAGACCACTCAAGTGC 4184
Db 1810 ACCAAAGAAATGGGCTTCAACTCGCGAGATCTTATTAAGCAACAGACCACTCAAGTGC 1869
Qy 4185 AAGGTAATTTGACATTCATTCCTCGCTTGACACCGACAGTCTGCCCGCTTCCGTTAGCTC 4244
Db 1870 AAGGTAATTTGACATTCATTCCTCGCTTGACACCGACAGTCTGCCCGCTTCCGTTAGCTC 1929
Qy 4245 ACAAGCTTACAGTACAGAGAGGTTCAAGAGCATCACCTCACTGACCTGCAATGGGAC 4304
Db 1930 ACAAGCTTACAGTACAGAGAGGTTCAAGAGCATCACCTCACTGACCTGCAATGGGAC 1989
Qy 4305 CAACATTTGCTGACAAAGAGAAATTTGGGGCTGCGAGAGAGCAAGCAAGCAAGATGATTA 4364
Db 1990 CAACATTTGCTGACAAAGAGAAATTTGGGGCTGCGAGAGAGCAAGCAAGCAAGATGATTA 2049
Qy 4365 CAGGCTTACATCCAGAAATTTTCTGTGGGGCGAAGAGGCTGGAGTACGTATGGGCTA 4424
Db 2050 CAGGCTTACATCCAGAAATTTTCTGTGGGGCGAAGAGGCTGGAGTACGTATGGGCTA 2109
Qy 4425 ACCATGAACAGTACAGAGTGTGGGCGAGAGTGGGACCAAGGCGACCCACATGATGGC 4484
Db 2110 ACCATGAACAGTACAGAGTGTGGGCGAGAGTGGGACCAAGGCGACCCACATGATGGC 2169
Qy 4485 CGCATGAGATCATCATCACTATTTATCATCGGCAATCCAGTCTCACTGTGCTGT 4544
Db 2170 CGCATGAGATCATCATCACTATTTATCATCGGCAATCCAGTCTCACTGTGCTGT 2229
Qy 4545 GTGTGTCTGCTCTTGGTATCTGTGTAGGCACTGATCATACAGCGTTTGCATGCCCAAAG 4604
Db 2230 GTGTGTCTGCTCTTGGTATCTGTGTAGGCACTGATCATACAGCGTTTGCATGCCCAAAG 2289
Qy 4605 CAAGAAGAGCTGCTGACGCAATAGCGGCTTGCACGAAAGCAAGCGTAAACCCACAGCAT 4664
Db 2290 CAAGAAGAGCTGCTGACGCAATAGCGGCTTGCACGAAAGCAAGCGTAAACCCACAGCAT 2349
Qy 4665 TAGCGGTTTGTGCTGATTCGCGCAACCAACGCTGAAACATTTGGAGAAACCTTGAACC 4724
Db 2350 TAGCGGTTTGTGCTGATTCGCGCAACCAACGCTGAAACATTTGGAGAAACCTTGAACC 2409
Qy 4725 ATCTGTGTTTAAACAAACCAACCGTTTCTGTGGGCAAGTGTGATTCCTGTGCAAGCGC 4784
Db 2410 ATCTGTGTTTAAACAAACCAACCGTTTCTGTGGGCAAGTGTGATTCCTGTGCAAGCGC 2469
Qy 4785 TTGTTATTTCTGTTCCGCTGCTTTCATGCTGCAAGCTTTTATTTAGGTTGCAAGCGCT 4844
Db 2470 TTGTTATTTCTGTTCCGCTGCTTTCATGCTGCAAGCTTTTATTTAGGTTGCAAGCGCT 2529
Qy 4845 GCCTGGGAAGTGAAGCGCTTTCGAACATGCAACACTGTGCAAAATGTTCCGGGGAATCC 4904
Db 2530 GCCTGGGAAGTGAAGCGCTTTCGAACATGCAACACTGTGCAAAATGTTCCGGGGAATCC 2589
Qy 4905 CGTAAAGGCGTGTGTCGAACCGGCAAGTTAACCGGCACTTAACCTGAGATCAAGGCTGC 4964
Db 2590 CGTAAAGGCGTGTGTCGAACCGGCAAGTTAACCGGCACTTAACCTGAGATCAAGGCTGC 2649
Qy 4965 TCTCATCGAAATTAACACTTCACTTAAAGAGTACGTGACCTGCAATTCACACAG 5024
Db 2650 TCTCATCGAAATTAACACTTCACTTAAAGAGTACGTGACCTGCAATTCACACAG 2709
Qy 5025 TCAATTCCTTCAACCAAGTTAAATGCTGCGGGTCCCTCGATGCAAGGATCCTCAAGG 5084
Db 2710 TCAATTCCTTCAACCAAGTTAAATGCTGCGGGTCCCTCGATGCAAGGATCCTCAAGG 2769
Qy 5085 CGGATTAACATCCCGGTTTGGGCTGTGATACCTTTCACTGAGGAGGCGCAAT 5144
Db 2770 CGGATTAACATCCCGGTTTGGGCTGTGATACCTTTCACTGAGGAGGCGCAAT 2829
Qy 5145 GCTTCTGTGACAGTGAAGAACACAACTGAGAGGCTGACGTGAGTTGCTTCAGACT 5204
Db 2830 GCTTCTGTGACAGTGAAGAACACAACTGAGAGGCTGACGTGAGTTGCTTCAGACT 2889
Qy 5205 GCACTATAGATCAAGAGTGCACATAAAGTTCAACAGCTGCTGAAAGTGGCTGC 5264
Db 2890 GCACTATAGATCAAGAGTGCACATAAAGTTCAACAGCTGCTGAAAGTGGCTGC 2949
Qy 5265 GTATAGTATCGGCAACCAACCGGCACTGATACGTTTGTCAATGGGCTGACGCGAG 5324
Db 2950 GTATAGTATCGGCAACCAACCGGCACTGATACGTTTGTCAATGGGCTGACGCGAG 3009
Qy 5325 GTTCTTCAAGGAGCTGAAGGATATAGAGGCGATATACGCGCTTTTCAACCTTTG 5384
Db 3010 GTTCTTCAAGGAGCTGAAGGATATAGAGGCGATATACGCGCTTTTCAACCTTTG 3069
Qy 5385 ACCATTAAGTGTGATCAGAAAGGCTGTTTAACTACAGCTTCCGTAATGGAG 5444
Db 3070 ACCATTAAGTGTGATCAGAAAGGCTGTTTAACTACAGCTTCCGTAATGGAG 3129
Qy 5445 CTATGAACAGAGAGGTTGGGATTAATCAACATCTGCTGATGCTACAGACATAG 5504
Db 3130 CTATGAACAGAGAGGTTGGGATTAATCAACATCTGCTGATGCTACAGACATAG 3189
Qy 5505 TAGCCGCACTGACATACGCGCTGTAAGCCTTGTGCAAGAAATTCACAGTCCCTTACA 5564
Db 3190 TAGCCGCACTGACATACGCGCTGTAAGCCTTGTGCAAGAAATTCACAGTCCCTTACA 3249
Qy 5565 CCAAGCAGTATCAGGGTATGAATGTGAAGAACTACAGACCAACCCCTGCAAGAAA 5624
Db 3250 CCAAGCAGTATCAGGGTATGAATGTGAAGAACTACAGACCAACCCCTGCAAGAAA 3309
Qy 5625 CAGACCATTTGATGTAATGAATGAGGCTGCGAGGCTGCACTGATGCTTACG 5684
Db 3310 CAGACCATTTGATGTAATGAATGAGGCTGCGAGGCTGCACTGATGCTTACG 3369
Qy 5685 GGCACATCCCTATCTGATGATGATCCTGATGAGCTTTTGTGATCATCAAAATCAC 5744
Db 3370 GGCACATCCCTATCTGATGATGATCCTGATGAGCTTTTGTGATCATCAAAATCAC 3429
Qy 5745 CAACAAATTTAGAAAGTTAGCTGCAACAGTACAGATGATTTATTTGCAAGCTTTGGTG 5804
Db 3430 CAACAAATTTAGAAAGTTAGCTGCAACAGTACAGATGATTTATTTGCAAGCTTTGGTG 3489
Qy 5805 GTTCTCTTAATTAAGTAAAGCTGACAGGAGGCAATGTCAGATTCACCTCCACT 5864
Db 3490 GTTCTCTTAATTAAGTAAAGCTGACAGGAGGCAATGTCAGATTCACCTCCACT 3549
Qy 5865 CCAAGCAGCTGTTTGAAGAGAGCAACACATGATGATGCTGAGGAGGAGATTAAC 5924
Db 3550 CCAAGCAGCTGTTTGAAGAGAGCAACACATGATGATGCTGAGGAGGAGATTAAC 3609
Qy 5925 TACATTTTACACATGAGCCCAAGCAAAATTTATGTTTGGCTATGCGGCAAGAGT 5984
Db 3610 TACATTTTACACATGAGCCCAAGCAAAATTTATGTTTGGCTATGCGGCAAGAGT 3669
```

5985 CCACCTGCAATGCTGAATGTAAACCCGCGCCGACATATTTGGAGAACACATTAAG 6044
3670 CCACTGCAATGCTGAATGTAAACCCGCGCCGACATATTTGGAGAACACATTAAG 3729
6045 TCGACCAAGAAATTCAGCGCGGAGCTTTCCAAAACCTTTGAGATGGCTGCTTGACGT 6104
3730 TCGACCAAGAAATTCAGCGCGGAGCTTTCCAAAACCTTTGAGATGGCTGCTTGACGT 3789
6105 TTGGGGGAGCATCCCTCATTTGTTGAGACTTATAGTGTGTGCTGACCTTATGC 6164
3790 TTGGGGGAGCATCCCTCATTTGTTGAGACTTATAGTGTGTGCTGACCTTATGC 3849
6165 TTATTAACACAGTATGATGACTGACCGCGGACACTGACATAGCGGTAAACCTGATGAC 6224
3850 TTATTAACACAGTATGATGACTGACCGCGGACACTGACATAGCGGTAAACCTGATGAC 3909
6225 TTCCGAGGAGCGTGTGCTGATTAATGCCACGCGCGCTTGACACTAAACCTGATGATTT 6284
3910 TTCCGAGGAGCGTGTGCTGATTAATGCCACGCGCGCTTGACACTAAACCTGATGATTT 3969
6285 CCGAGGAGGACAGTGCATTAATGCTGTGAGTGCATTAATGCTGTATATACACTACAT 6344
3970 CCGAGGAGGACAGTGCATTAATGCTGTGAGTGCATTAATGCTGTATATACACTACAT 4029
6345 ATTAACACACTATATATCTTTATGAGACTGACTATGGGTCTTAATATACACTACACA 6404
4030 ATTAACACACTATATATCTTTATGAGACTGACTATGGGTCTTAATATACACTACACA 4089
6405 TATTTACTTAAAAACACTATACACTTATTAATTTCTTTATTAATTTTCTTTGTTT 6464
4090 TATTTACTTAAAAACACTATACACTTATTAATTTCTTTATTAATTTTCTTTGTTT 4149
6465 T 6465
4150 T 4150

RESULT 4
US-09-507-362-103
Sequence 103, Application US/09507362
Publication No. US20030096397A1
GENERAL INFORMATION:
APPLICANT: Dubensky Jr., Thomas W.
Polo, John W.
Bell, Barbara A.
Schlesinger, Sondra
Dryga, Sergey A.
Frolov, Ilya
TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
SYNTHESIS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESS: Seed Intellectual Property Law Group PLLC
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,362
FILING DATE: 18-Feb-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McMaetere, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 930049.457D6 /1196.011

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 11740 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 103:
US-09-507-362-103
Query Match 36.7%; Score 2380.4; DB 11; Length 11740;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3693; Conservative 0; Mismatches 1986; Indels 41; Gaps 8;
580 ACTGACGATGAAAGCGGAGCGTATATTTCTCATCGGAAACAGGCCAAGTCACTTC 639
5746 ACTGACTAACCGGGGTAGTGGTATCATATTTTCACGACAGGCCCTGGGCACTTGC 5805
640 AACGAATTCAGTACGTAATGTAACACAAACCTATATTTGATGGGCCCTGCATG 699
5806 AAAAGAAAGTCCGTTCTGAGAACCAAGCTTACGAAACGACCTTGAGGCAATGCTGG 5865
700 AGAAGTATTACGCCCGCGCTCGATCTCGAAAGAGAAATGTTACAGAAAGAACTGC 759
5866 AAGCAATTCATGCCCCGCTGCTGACAGCTGGAAGAGAAACCACTCAATCAGGTACC 5925
760 AATTATGCGCTCTGAAAGAAATAGAAAGAGATTCATCAAGAAAGTAAATATGA 819
5926 AGATGATGCCACGAGCAACAAAGTATGATGACAGTCTGTAAGTAAAGTAAATATGAGA 5985
820 AAGCAATTCAGCGGAGGAGTCAATTTCTGATGGGACATATCTATCAAGAAAGTGA 879
5986 AAGCATTAACCACTGAGGAGTACTGTCAGACTCAAGCTATTAATCTTCCACAG--- 6042
880 ATCTGTGAGTGTATACAGATCAATTAATCTGTAACCAATCTACTCTGCAACGGTAATTA 939
6043 ATCAAGCCAGATGCTATTAAGATCACTATCCGAAACCTATGTAATCTCAGTACGATCCGG 6102
940 ACAGTTTACATCTGACAGAGTGGCGGTTAAACGTGCACTTAATTCAGAAAGAAATT 999
6103 CGAATCTACTCGGATCCAGATGCTGCTGATGCTGCTGATGCTGCTGATGCTGCTGATGCT 6162
1000 ACCCTACAGTACGAGTATTTGTAATTAAGATGAATAGATGCGTATTTGACATGCTGG 1059
6163 ATCCGACAGTACATCTTATGATTAAGATTAAGATGATGATGATGATGATGATGATGATG 6222
1060 ACAGGAGATCGTGTCTGATATACAGCACTTTTTCGCGCTAACTGAGAAAGTACC 1119
6223 ACAGGAGATCGTGTCTGATATACAGCACTTTTTCGCGCTAACTGAGAAAGTACC 6282
1120 CAAGAAAGATAGCTATTTTTCGAGCCAGATTAAGATGACGCGCTGATGCTGATGATGATG 1179
6283 CGAAGAAAGATAGCTATTTTTCGAGCCAGATTAAGATGACGCGCTGATGCTGATGATGATG 6342
1180 ATACATTAAGAAATGATTTGCTGAGTATTAAGAAAGATGCAAGTATTCACCAATGTC 1239
6343 ACAGCTTAAGAAATGATTTGCTGAGTATTAAGAAAGATGCAAGTATTCACCAATGTC 6402
1240 GAGAAATACCTGCTTATGATTTGAGGAGGATTAATGTTGATTTTCAAGAAATTCGAT 1299
6403 GTGAACTGCCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6462
1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
6463 GTAATGAGAGTATTTGAGAGAGTTGCTGAGAGCCATTAAGATTAAGATGATGATGATGAT 6522
1360 TTACGCAATATGTAAGAAAGTGAAGAGGCGGAAAGCAGCAGCATTTGTTGCAATATCTC 1419
6523 TCACCGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6582
1420 ATATATCTAAACCGTTGAGAGATTAACATGATGATGATGATGATGATGATGATGATGATGAT 1479

```
Db 6583 ATATATTTGGTCCCATTTCCAGAGAGTGGCTATAGATAGATTGTCATGGAACATGAAAAG 6642
Qy 1480 ATGTCAAAAGTTACTCTCCGGCAGAAACATACAGAGAGCGGCTTAAGTGCAGTTATTC 1539
Db 6643 ACGTGAAAGTTACACACAGGCGACGAAACACACAGAAAGAAAGCCAAAGTACAAAGTATAC 6702
Qy 1540 AGGCTGAGATCCCTTGTACCGCTTACCTTTGCGGGATCCATCGGGAATTAAGTCGTA 1599
Db 6703 AAGCCGAGAACCCCTGGCCACTCTTACTTATCGGGATTACCGGGAATTAAGTCGTA 6762
Qy 1600 GACTGAATGCGGTCTTTCGCAATATCCACTCTTTCGACATGTACGCGGAAGATT 1659
Db 6763 GGGCTTACGGCGCTTCTGCTTCAAAATGACACGCTTTTTCGACATGTCCGCGAGATT 6822
Qy 1660 TTGATGCGATTATTTGCTGAACATTTCCACACGGCGACCCAGTATGGAACGACATCG 1719
Db 6823 TTGATGCAATCATATGACGAAACATTTCAAGCAAGCGACCCGCTACTGGAGACGATATTCG 6882
Qy 1720 CGTGGTTGATTAAGGAGAGAGACGCTATCGCATTTGGCGGCTTATGATCCCTGAGG 1779
Db 6883 CATATTCGACAAAAGCCAGACGACGCTATGGCGTTAACGGTCTGATGATCTTGAGAG 6942
Qy 1780 ACTTAGGTGTCAGCAACCGCTTATGATTTGATAGAGCGCGCTTCGCAATATCAAT 1839
Db 6943 ACCTGGGTGGATCAACACTACTGACTTGTATCGATGGCGCTTTGGAGAAATATCAT 7002
Qy 1840 CTGTGCACCTTACTCAACGAAACGAGTTTAAATTTGGTCCATATGAAATCCGATATGT 1899
Db 7003 CCACCCATCTACCTACCGGATCTCGTTTAAATTCGGGGCGATATCAATCCGGAATGT 7062
Qy 1900 TCTTAAGCTGTTGTGCAACACACTAGTCAATATCATGATTTGTAGAGAGTACTAGTG 1959
Db 7063 TCTTCAACATTTTGTGACACAGATTGTAATGTCGTTATGCGCAGAGAGTACTAGAG 7122
Qy 1960 AACGGTTAACCAAGCTGACGCGCGGCTCTATCGCGAGAGATTAATAGTGCATGTG 2019
Db 7123 AGGGGCTTAAACGTCAGATGTGCAAGGTTCA TTGGGAGACGACAACTCATATCATGAG 7182
Qy 2020 TCGTCTCCGACACTTATAGCGGAGAGATGCGCACCTGGCTGAACATGGAAGTAA 2079
Db 7183 TAGATCTGACAAAGAAATGGCTGAGAGGTGCGCACCTGGCTCAACATGAGAGTTTAA 7242
Qy 2080 TTATTTAGAGATTAATTTGATTAAGAACCCCTACTCTGCGGGGATTAATCCGTGAG 2139
Db 7243 TCATCGCGAGATATGAGTGAAGACCACTTACTTCTGCGGGGATTAATCTTGCAAG 7302
Qy 2140 ACCAGATTAACAGGACGCTGACAGATGCGAGACCTCTAATAAGGCTTTTAAAGCTTG 2199
Db 7303 ATTCGGTTACTTCCACAGCGGCGCGGCGATCCCTGAATAAGGCTGTTTAAAGTTGG 7362
Qy 2200 GAAAAACATTGCGCATGATGATACCAAGATCGACCGCGCGGCACTGATATG 2259
Db 7363 GTAAACCGCTCCCGCGAGACAGACGAAAGCAGAAAGCGGCTCTGTCTAATAGT 7422
Qy 2260 AAGCAATGCGATGGAACAGATTGGAATTACGAGAGTTAGTAAGACCGGTAGATCCA 2319
Db 7423 AAACAAAGGCGTGTGTTTGAAGTAAATTAACAGGCACTTTAGCAGTGCCTGACGACCC 7482
Qy 2320 GATACGAGATCATCTGCGAGGCTGTATCATCATGCTCTGTCACGTTAGCCGAAAGCG 2379
Db 7483 GGTATGAGTGAACAATATTAACCTGTCTTACTGAGATTTAGAACTTTTGGCCAGAGCA 7542
Qy 2380 TTAAAGACTTCAAGAGCAATAAGAGGAGCCCAATACCTCTTACGCTGACCTTAATAGG 2439
Db 7543 AAAGAGCATTCCAAGCATCAGAGGGGAAATAAAGCATCTCTAAGGTGCTCTTAATAAGT 7602
Qy 2440 TGAAGTATGAGACGACGACCTAC-----CCACCGGAGAGATGTTCCATACC 2486
Db 7603 CAGCATATGATTTCAATCTGACTAATACTAACAAACACACCACTAGTAATAGAGATTTCT 7662
Qy 2487 CTCAGCTGAACCTTTCACAGTTTACCTTACCAATTCGATGGCTTACCGAGATCAAAACC 2546
Db 7663 TTAACATGTCTGGCCCGCGCCCTTCCGCGCCCACTGCCATGTGAGGCGCGAGAA 7722
Qy 2547 CTCTAGAGCCGCCCTGAGAGCCGTTTCCGCCCGCGTGGCTGCATTAATGAGATCTTA 2606
Db 7723 GAGAGCAGCGCGCCCGATGCTTCCCGCAAGGGCTGGCTTCTCAATCTCAGCATCTGA 7782
Qy 2607 GAGAGTCAATATGCACTTGACTTTCAACAAGATCACTTAATCCGCCCGCAGTCCA- 2665
Db 7783 CACAGCCGTCATGCGCTTATGATGAGCAGGCACTAGACCTCAACCCCAAGTCCAC 7842
Qy 2666 -----CCGCCAAAGAAAGAAAGATGCTCTTAAGCCAAAACCTTACCTTAAAGA 2720
Db 7843 GCCCGCACCGCCCGAGAAAGCAGCGCCCAAGCAACCCAGAGGCCAGAAACAA 7902
Qy 2721 AGAAGCAGCAAGCCAGAGAGCAAGAAC-----GCAAGCTTAACAGGGAAGCAACGTA 2777
Db 7903 AAACGACGAGGAAAGAAAGAAAGCAACCTGCAAAACCCAAACCCGAAAGAGACGCGCA 7962
Qy 2778 TGTGTATGAAGTTGAGTCCGACAGACATTTCCGATCATG-----CTGAACGCCAAGTGA 2834
Db 7963 TGGCACTTAAGTTGAGGCGCGACAGATTTGTCACCTCAAGAACGAGAGAGATGTCA 8022
Qy 2835 ATGATATGCTGCTGCTTTCGAGAGAGCTGATGAACCACTCCACTTGAAGAAAA 2894
Db 8023 TCGGGCAGCAGTGGCCATGGAAGAGAAAGTAAATGAACCTCTGCACTGTAAGAAACA 8082
Qy 2895 TTGATATAGCAATTAAGCGCGCTGAAATTTGAAGAGCTTACATGTAGACTTGAGAT 2954
Db 8083 TCGACCACTGTGCTATCAAACTCAAAATTTACCAAGTGTACATGACATACAGCATGAGT 8142
Qy 2955 ACCGCGAGCTTCCCGAGAACATGAAATCAGACACGCTGCAATCACCGAGCGACAAACAC 3014
Db 8143 TCCGACAGTTGCGAGTCAATAGAGAGATGAGGCAATTCACCTTACACAGTGAACACCCG 8202
Qy 3015 CGGGCTTCAACACTGGCACACCGCGGAGTCCAGATGTGAATGGAGATTTACCTGATAC 3074
Db 8203 AAGGATTTCTAATCTGGCACACGAGCGGTGAGTATGTGAGATGATTTTACCATCC 8262
Qy 3075 CAGAGAGTGGCGCGGGAAGGAGACAGCGGAAAGACCGATCCGTGCAACAGAGGCGAG 3134
Db 8263 CTGCGGAGTAGAGGACAGAGAGACGCGGTGCTGCAATAGATATCTCCGTTGGG 8322
Qy 3135 TTGCGCTATTTGTTAGAGGTGCAAAATGAGGCGACGCTTACGCGCTTTCAGTGTCA 3194
Db 8323 TTGTGCGGATGATCTCTGCGTGGCGCTGATGAAGAAACAGCACTGCCCTTTCGCTGTCA 8382
Qy 3195 CTTGGAACAGAAAGGGGTGACATTAAGGATATCCCGGAAGTTTGAACGCTGT--- 3251
Db 8383 CTTGGAATAGTAAAGGGAACATTAAGACGACCCCGGAAGGAGACAGAAAGTGTCCG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGGGTGCTTGAATGTCAGTTCCTCAATGGAACAA 3305
Db 8443 CAGCACTACTGTGTCAGGCAATGTGTTGCTCGGAATGTGACCTTCCATGGAACCGCC 8502
Qy 3306 CACCGGTGCTATTTCACTGACCCAGAAAGCAACCTTGAACCTGCTGGAAGAGAGATG 3365
Db 8503 CGCCCACTGTATACCCCGGAACCTTCCAGAACCTTCCAGAACCTTGAAGAGAGATG 8562
Qy 3366 ACAATCCAAATTAACGACACCTGCTGAGAGAGCTTTGAATGTCCATCAC---GCCGCG 3422
Db 8563 ACCATGAGGCTTACGATACCTCTCAATCCATATGTCGCTCGGATCTCTGCGAGAA 8622
Qy 3423 CCAACGAAAGCAATTAACGATGATTAACACTGACAGTCCCTTACCTGAGGCTTTCGCGCT 3482
Db 8623 GCAAAAGAACGTCATTTGAGACCTTACCTGACAGCCCTTACTTGGGAGCATGCTGT 8682
Qy 3483 ATTGACAGACCTCAAGCGCTGTTTACAGCCCAATTAATTTGAGAACTGTGAGAGCAAT 3542
Db 8683 ACTGCCACATATCTGATCGCTGCTTCAAGCCCTTTAAGATCGAGAGGTCTGGGAGAG 8742
Qy 3543 CTGATGATGATGATGATTAAGATCCAGGTCTGCGACAAATTCGCTTCAATCAGGAGCA 3602
Db 8743 CGAGAGTAAACCAATACGATACAGACTTCCCGCAGTTTGGATTCGACCAAAAGCGAG 8802
```

QY 3603 CTGGGATGTCACAAAATTCGGTTACATGTCCTTTGACACGACCACATGATCAAGANG 3662
DB 8803 CAGCAAGGCAAAACAGTACCGCTACATGCTTAAACAGAGTACACCGTTAAAGAG 8862
QY 3663 ACAGATGAGAGAAATAGCTATCAGACATCTGCACCTCGCTGCTTGGCCAAAG 3722
DB 8863 GCACCATGAGATCAATCAAGATTAGCACCTGAGGACCGTGTAAAGCGCTTGTACAAAG 8922
QY 3723 GGATCTCTGTTAGCTCAATGTCCTCGAGTGACAGTAAACGCTGATTCACAGACG 3782
DB 8923 GATCTTCTCTCGCAAAATGCCCCCTCCAGGAGACAGGTAAACGTTAGCATGTAAGTA 8982
QY 3783 GAGCATCTGAGAAATTCATGACACCGTGGAGAAAAAGATCAGAGGAAATTTGCTGGTAG 3842
DB 8983 GCAACTAGCAACGTCAATGTAACATGCGCCCGCAAGTAAACCAAAATTCGTGGAGCGG 9042
QY 3843 AGGAGTACTTGTCCACCGCTCCATGGAAAGCTGGTAAAGTGCCACGTTTACGATCACT 3902
DB 9043 AAAAATATGATCTACTCCGTTACCGGTAAAAAATTCCTTGACACAGTGTACGACCGTC 9102
QY 3903 TGAAGAGACGCTCGCGGGGTACATTAACATGACAGGCGCCACACCGCTATAGT 3962
DB 9103 TGAAGAAACAACTGACAGGCTACATCACTATGACAGGCGCAGACCGCATATATCAT 9162
QY 3963 CCTATCTGAGAGAGCGTCAAGCGAAGTGTACATTTAAACCACTTCTGGCAAGACGTC 4022
DB 9163 CCTACTGGAAGAAATCATCAAGGAAAGTTTACGAAACCCGCCACTTGGAAAGAACATTA 9222
QY 4023 CCTACGAATGTAGTGTGCGCATACACACACAGTATCTGACACCGGAAAGAAATGA 4082
DB 9223 CGTATGAGTGAAGTGGCGGACATCAAGACCGGAAACGTTTTCGACCGGCAACCAATCA 9282
QY 4083 ACGGCTGACATAAAGCAAAACAGTGCATTGCTAAGAGGACCAAAACGAATGGGCT 4142
DB 9283 CTGCTTGACACCGCATCAAGCAGTGCCTGCTTAAAGAGGACCAACGAAGTGGGCT 9342
QY 4143 TCAACTGCGCGGATCTTATTAGGACACAGACCACTCAGTCAAGTAAATTTGCATTC 4202
DB 9343 TCAACTCAGCGGACTTGATACAGATGACGACACACGCGCCCAAGGAAATTTGCATTTGC 9402
QY 4203 CATTCGCGTTGACACCGACAGTGTCCCGGTTCCGTTAGTCAACGCGCTTACAGTCA 4262
DB 9403 CTTCAGATGTATCCGAGTACTGTCATGTGCTCCGTTGCCACCGCGCAATGTAAATAC 9462
QY 4263 AGTGTTCAAAGGATCACTCCCTCAACCTGACCTGCAATGCCAACATTTGCTGACAA 4322
DB 9463 ATGGCTTTAAACATATCAGCTCCATTTAGATACAGACCACTTGACATTTGCTCAC 9522
QY 4323 GAAATTTGGGCGTGGACAGACAGCAACAGAGATGATTAACAGGCTCTACATCCAGGA 4382
DB 9523 GGAATCTAGGGGCAAAACCCGGAACCAACCACTGAATGATGCTGGAAGACCGGTCA 9582
QY 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTACGTATGGGTTAACCATGAACAGTCAAG 4442
DB 9583 ACTTCAACGTTGACCGAGATGGCTTGAATACATATGGGAAATCATAGGCAAGTGA 9642
QY 4443 TCTGGGCGCCAGAGTGGGACACAGGCAACCAATGATGGCCGATAGATCATATCC 4502
DB 9643 TCTATGCCCAAGATCAGACACAGAGAACCTTCAACGATGGCCACACCAAAATAGTAC 9702
QY 4503 ACTATTTATCATCGGATCCAGTCTACACTGTCAATTTGCTGTGTGTGCTGTGCTA 4562
DB 9703 ATTACTACATCGCATCTCTGTGTACACATCTTAAGCCGTGCACTACGATCCGTGCG 9762
QY 4563 TCTGTAGAGGACGTCATCATCAGACGTTGATCGCAAGAACAGAGAGATGCTCTGA 4622
DB 9763 TGATGATGGCGTAATCTGTGTGATGTATGTGCTGTAAAGCGCGGTGAGTGGCTGA 9822
QY 4623 CGCCATACGCGCTTGACCGAAACGTAACGTAACCAAGCATTTAGCGGTTTGTGTGCTGA 4682
DB 9823 CGCCATAGCGCTTGCGCCCAACCGCGTAATCCCAATTCGCTGACCTTGTGTGCTGCG 9882

QY 4683 TTCCGCCAACCAACGCTGAAACATTTGGAGAACTTTGAACCATGTGTGTTTAAACAAC 4742
DB 9883 TTAGTGGGCCCAATGCTGAAGAAAGTTTACCCAGACCATAGATTACTTGTGTGCAACATC 9942
QY 4743 AACCGTTTCTTGGGACAGTTGTGATCTCTCTGGGACGCGCTGTATTTCTGTTCGCT 4802
DB 9943 AGCGTTCTTCTGGGTCCAGTTGTGATACCTTTGGCGCGCTTTCATGCTTCAATGCCCT 10002
QY 4803 GCTTTGATGCTGATGCTTTTATTTGTTGTGAGGCGTGTGCTGGGAAAGTAGACG 4862
DB 10003 GCTGCTCTGCTGCTGCTTTTATTTAGTGTGCGGCGCTTACTGGCGAAGTAGACG 10062
QY 4863 CCTTGAACATGCGACCATGCTGCAATGTTCCGGGATCCGTAATAGGCGTGTGCTG 4922
DB 10063 CCTAGAAACATGCAACCATCTTCCAAATGTGCCACAGATACGTAATAGGCACTGTG 10122
QY 4923 AACCGAGGTTACGCGCCACTTAACCTGAGATCAAGGTGTCTCATGGAAATTAACAC 4982
DB 10123 AAAGGCGAGGTTAGCCCGCTCAATTTGGAGATCACTGTATGCTCTCGAGGTTTTC 10182
QY 4983 CTTCACTAACAGAGTATGCTGACCTGCAAAATTCACACAGTCAATTCCTTACCCACAG 5042
DB 10183 CTTCAACCAACCAAGATACATTAACCTGCAAAATTCACACAGTGTGCTCTCCCAAAA 10242
QY 5043 TTAATGCTGCGGCTCCCTGAGTGAAGGATCCTCAAGGCGGATTAACATGCGCG 5102
DB 10243 TCAATGCTGCGGCTCCTTGAATGTGACGCGCGCTGCTCAATGACATATACCTGCAAG 10302
QY 5103 TTTTGGCGGTGTACCTTTCAATGTGGGAGGCGCAATGCTTCTGTGACAGTGA 5162
DB 10303 TCTTCGAGGGGTCTACCCCTTATATGTGGGAGAGGCGCATATTTTTCGACAGTGA 10362
QY 5163 ACACAACTAGAGGAGGCGGATGAGTTCGATTCGTCAGACTATATAGTACAGGAG 5222
DB 10363 ACAGCCAGATGAGGAGCGGATCCTCAATTTGTACAGATTTGGCGCTTACCAACGCGC 10422
QY 5223 TCGCACTAAAGTTTCAACAGCTCTCTGAAAGTCCGCTGCTATAGTATACGGAACA 5282
DB 10423 AGCGGATTAAGTGTACACATGCGCGCATGAAAGTANAGATCTGCTATGTGTACCGGAACA 10482
QY 5283 CCACGCGCACCTGATACGTTTGTCAATGGCGTACGCGCAGGTTCTCAGGGAACCTGA 5342
DB 10483 CTACAGATTTCTGTAGATGTGTACGTGAACGAGTCAACACGAAACGCTTAAAGCTTGA 10542
QY 5343 AGGTATAGCAGGCGCATATCAGCCGCTTTTACACCTTTGACATTAAGTCTGTATCA 5402
DB 10543 AAGTCATAGCTGACCAATTTCAACATCTTTTACGCAATTCGATCAATTAAGTCTGTATCC 10602
QY 5403 GAAAGGGGCTTGTTAACAACATGCACTTCCCTGAGTATGAGCTATGAACACAGAGGCT 5462
DB 10603 ATCGCGGCTGTGTACACATATGATCTTCCGGAATTTGAGCGATGAAACCAAGAGCT 10662
QY 5463 TCGGCGATTTCAAGCATCTCTGCTGATGTACAGACATATAGTACCGGACCTGACATAC 5522
DB 10663 TTGAGACATTTCAAGTACCTCTTACTAGCAAGATATTCATGCGCAGACACATTA 10722
QY 5523 GCGTGTGAAGCCTTCTGTCAAGAACTCAACGTCCTCTACACCAAGCAATATCAGGCT 5582
DB 10723 GGTACTCAAGCCTTTCGCGCAAGAACGTGATGTCCGTACACAGGCGCTCATCAGAT 10782
QY 5583 ATGAATGTGGAAGAACAACTCAGAGAGACCCCTGCAAGAAACAGACCAATTTGATGTA 5642
DB 10783 TTGAGATGTGGAAGAAACATCAGGCGCGCCATCTGCAAGAAACCGCACCTTTCCGGGTGA 10842
QY 5643 AAATTAAGTGAAGCCTCTGCGAGCGTCTAATCTGTGTTACGGGACATCCCTATCTGA 5702
DB 10843 AGATTGCAATAATCCGCTCGAGCGGCTGAGCTGTTCATACAGGGAACATTTCCATTTCTA 10902
QY 5703 TTGACATCCCTGATGAGGCTTTTGTGATCATATAGATACCAACAAATTTTAAAGTGA 5762
DB 10903 TTGACATCCCGAAGGCTGCTTTATATCAGACATCAAGTACCACTGTCTCAACAGTGA 10962
QY 5763 GCTGCAAGTAGACAGACTGATTTATTTCTGACAGACTTGTGTGTTCTTAACATTAAGT 5822

Qy 1194 GTATTGGCTGCACTACTATAAAGAAATTGCACGTTACCCAAATCGAATAATTCCTGTC 1253
Db 6129 GTGTAGCGGCGCGCACCAAGAAAGAACTGCACAGCAAAATCGAAGAACTACCAACC 6188
Qy 1254 TTAAATTGGGGGCGCATTTAATGTTATGTTTCAAGAAATACGCATGCAATGATAGTAC 1313
Db 6189 ATGACTGGGAGGTTCACACGTGAGTGTTCMAAGCGCTATGCTCCCGAATAAT 6248
Qy 1314 TGGGATACCTTTCGGATTAACCTTAATTCGGCTAATACAGAGAGGTTACCGAATATG 1373
Db 6249 TGGGAAGATATGCTAAACACCTTAATCCGATTAACCACTGAGAACATCACTACCTATG 6308
Qy 1374 ACAAGCTGAAAGGCGGAAAGCAGACAGATGTTGGAAATACATTAATCTTAAACCG 1433
Db 6309 ACCAAATGAAAGCGCGAAAGCTGCTGCTTGTTCGTAAGACCAACACTGTTCCG 6368
Qy 1434 TTGACGAGATACCAATGATCAATTCGTATGATCTTAAGAGATGTCAAAGTTACT 1493
Db 6369 CTGACGAGGTTCCCATGAGACGATTCACGCTGACATGMAACGAGATGTCAAAGTCACT 6428
Qy 1494 CCGGCAAGAAACATACAGAGAGGCGCTAAGGTGACGTTATTCAGGCTGCAGATCC 1553
Db 6429 CCAGGAGGAAACACACAGAGGAAAGAACCAAGTCAGGTATTAACAGCGAGCA 6488
Qy 1554 CTGCTACCGCTTACCTTTGCGGATCCATCGGAAATTAATCCGTAGACTGATGCGGTG 1613
Db 6489 TTGGGACCGGTACTCTGTGGGATCCACAGGGAATTAAGTAAGAGACTAATCTGTG 6548
Qy 1614 CTTCTGCCAAATATTCATCTCTCTTGCACATGTCAGCGGAAATTTGATGCGATTA 1673
Db 6549 TTACCCCTTAACGTACACATTTGTTATGTCGCGGAAAGCTTGAACGCGATCATC 6608
Qy 1674 GCTGAACATTTCCACAGCGGCGACCATTAATGMAAGGACATGCGCTGTTATAA 1733
Db 6609 GCTCTCACTTCCACAGAGAACCCGCTTCTAGAGCGGACATTCATTCACAAA 6668
Qy 1734 AGCGAAGACGACGCTATCGCCATTTTCGCGTATGATCTCTTGAGACTTGAAGTGCAC 1793
Db 6669 AGCGAGACGACTCTGTGCTTACAGGTTAATGATCTCGAAGACTAGGGGTGAT 6728
Qy 1794 CAACCGCTTATGATTTATAGAGGCGGCTTCGCAATATCACTCTGTGACCTTACT 1853
Db 6729 CAGTACTGTGGAATTTATCGAGGCGACCTTGGGAAATATCCAGCTGTACCTACA 6788
Qy 1854 ACAGGAAGAGGTTAAATTTGTGTCATGATGAATCCGATATCTTAAACGCTGTT 1913
Db 6789 ACTGCGACGCTTCAAGTTGGAAGCTATGATGAATCGGCGATGTTCTGACTTGT 6848
Qy 1914 GTCAACACACTAGTCAATATCATGATGCTAGCAGAGTACTAGTGAACGTTAACAG 1973
Db 6849 ATTAACTGTTTGAACATCACTACATAGCAGAGGTAAGTGAACAGACTCACTGAC 6908
Qy 1974 TCAAGCTGCGGCGCTTATCGGCGAGATTAATAGTCATGTGTCTCCGACAC 2033
Db 6909 TCCGCTGTGCGGCTTATCGGCGAGCAAACTGTTACGAGATATCTCCGACAG 6968
Qy 2034 TTGAAGGCGGAGATGCGGCGACCTTGCTGAACATGAGAAATTAATTAATGATGCA 2093
Db 6969 CTGAAGGCGGAGGTGCGGCTGTGGGTCAACATGAGGTGAATATATGACCTCTTC 7028
Qy 2094 ATTGGATCAAGACACCTTACTTGTGGGGGATTTATCTGTGACAGATTAACAGG 2153
Db 7029 ATGGCGGAAAGACCCCATATTTTGTGGGGATTCATATGTTTGAACAGGCTCACAG 7088
Qy 2154 ACAGCTGCAAGTGCAGACCTCTTAAAGAGCTTTTAAGCTTGAAGAAACATATGCA 2213
Db 7089 ACCGCTCGCGTGTTCAGACCACTTAAGCGCTGTTCAAGTGGTAAGCGCTAAC 7148
Qy 2214 GTCAATGATACCAAGACTGCGACCGCGCGGCACTGCAATGAAAGCAATGCGATG 2273
Db 7149 GCTGAAGCAAGCAAGAGAGAGGCGAGCAAGAGCACTGAAGAGGATTAAGCAAGTGG 7208
Qy 2274 AACGAATTTGAATTAAGGAGATTAAGTGAAGCGCTGATTAATTCAGATACGATCA 2333

Db 7209 TTCCGAGCAGGCTTGGGGCGCAACTGAGGTTGCACTAATCTAGATATAGTAGAG 7268
Qy 2334 CTGCGAGGCTGATCATCACTCTCTCTCCACGTTTACCGGAAAGGTTAAAGACTTCAAG 2393
Db 7269 GCGTCAAAAGATCTCTATAGCCATGACACCTTGGGAGGAGCATTAAGCGCTTTAAG 7328
Qy 2394 AGCATTAAGAGGAGCGCAATCACTCTTACGAGTGAACCTTAATATAGTAGAGTAGACA 2453
Db 7329 AATTTAGAGAGACCTGTTATACACTCTTACGCGGCTCTTAATGTTGTGCTTAATAC 7388
Qy 2454 CGACCTTACCAACCGGCGAAATGTTTCCATACCTCACTGAATCTTCCACAGTTTAC 2513
Db 7389 GAATCTGATTAATAGCGCACTATTATAGCAACATGAATTAATCACTCCACGAAAGT 7448
Qy 2514 CTACAAATCCGATGCTTACCGAGATCCAAACCTCTCT-----AGGCGCGCTG 2563
Db 7449 ACAGCGCGGATGCGCGCGCGCGCGCGCGCTTGGCGCTTGGACGACCTCCG 7508
Qy 2564 AGCGCTTTCGCGCGCG-----CTGGCTGCTCAATGGAATGTTAGAGGTGATAG 2618
Db 7509 TGGCTCCCTGCTGCTCCGACTTCCAGGCTCAGCAGATCAGCACTCATCAGCGCTTA 7568
Qy 2619 TCAACTGACTTTCAACCAAGATCACTTAATCCGCGCCAGTTCACCGCAAGAA 2678
Db 7569 ATGGCTGACATGAGACAGAACGAAATGCTCTGCTGATAGGCTCCAAACCAAGAA 7628
Qy 2679 AGAAGAGTCTCTTACGCAAACTTACTACGCTTAAAGAGAGAGACAGCAAGCAAG- 2737
Db 7629 AGAAGACCAACCAACCAAGCGGAAACGAGCCCAAGATCAACGAAAGAGCGAGC 7688
Qy 2738 -----AGAGGAAACGCAAGC-----TAAACGAGGAAAGCAGAC 2774
Db 7689 AGCAAAAGAAAGAAAGACAGAGCGACAGCAAGAAAGAAAGAAAGAAAGAA 7748
Qy 2775 GTATGTATGATGAGTTGAGTGCAGCAAGATTTCCGATCATGCTGAACGCGCAAGTGA 2834
Db 7749 GAATGTGATGAAGTTGAATGATGCTGATCTTGAGAGTCAACAGAGAAAGTGA 7808
Qy 2835 ATGATATGCTGCTGCTTGTGCGAGAGAGCTGATGAACCACTCCAGTTGAAGAA 2894
Db 7809 CTGGTACGCTGCTGTGTGGCGACAAAGTCAATGAACCTGCCACCTGAAGAGTCA 7868
Qy 2895 TTGAATAGACCAATTTAGCGGCGCGGAAATTAAGAGAGCTGAGATGAACCTTGAAGT 2954
Db 7869 TGAACAGCGGACCTGCGCAAGCTAGCTTCAAGAAATCAAGCAAGTATGACCTTGA 7928
Qy 2955 ACAGGACGTTCCCGAGAACATGAATGACACAGCTGCACTACAGCAAGCAAGCAAG 3014
Db 7929 GTGCCAGATACAGATTCATGAGTGCATGCTCAAAATGATACAGATGAAGAGCCG 7988
Qy 3015 CGGCTTACACTGCGACCAAGCGCGCATGTCAGATATGAATGAGAGATTTACCGTAC 3074
Db 7989 AGGCACTATTAATGCGCACACCGGCGCTTCAATGACAGCGAGTGAAGTCACTAT 8048
Qy 3075 CGAGAGAGTGGGCGGAAAGCGGAGCAAGCGAATCCCTGGAACAGAGGCAAG 3134
Db 8049 CGACAGGCGGCGCAACCGGAGACAGTGCAGCGCCATCTTGAACAAAGGAGAGG 8108
Qy 3135 TTGTGCTATTTGTTCTAGAGAGTGCAAATGAGGCGACGCTACGCGCTTTCAGTGTCA 3194
Db 8109 TAGTGTGATGCTCTGCGCGGCGCAACGAGGCTCACGACAGCACTGTGCTGTCA 8168
Qy 3195 CTGGAACCAAGAGGCTGACATTAAGGATATCCCGGAAAGTTCGACCGTGT-- 3251
Db 8169 CTTGGAACCAAGATATGTTGAC---TAGATGACCCCGAGAGGCTCCGAAAGAGTGTCCG 8225
Qy 3252 ---CACTAGTTACAGCGCTATGCTGCTTGTGATGATGATCACTTCCAGCAAGAAC-- 3306
Db 8226 CCGCGTATTAAGTCAAGTGTGTCTTGGCAATGATCACTTCCGCTTCCAGGCC 8285
Qy 3307 -----ACCGTGTATTAAGTCAAGTGCAGAGAGCAAGCAAGTGCCTGCAAGAGA 3359

D8 8286 CGTGTACCTTGTCTGATATAAACAACGAGAGCCACACTACGATGCTCGAGATA 8345
QY 3360 ACCGTGCACATCCAAATTACGACACGCTGTGGAAGAGTTGTAATGTCCA---TCAC 3416
Db 8346 ACGGTGATAGCCGAGGGATCTACGACCTCTTCAGGACGCTTGACGTGCGAAACGGA 8405
QY 3417 GCCGGCCAAAAGAAAGATTACCGATGACTTCACACTG-----ACGATCCCTAAC 3467
Db 8406 CAGACACCGGCGGAGGTGTGCGAACACTTCAAGTGTATTAAGCTACACGCCCTTACA 8465
QY 3468 TGGGGTTCTGCGCTTATTCAGACACTCAACGCCGTGTTTCAGCCCAATAAATTAGA 3527
Db 8466 TCGGTACTGCGCGACTGCGAGACAGGGCACTGTGTCATAGCCCGTACACTTAAG 8525
QY 3528 ACGTGTGGACGATCTGATGATGATGATGATTAATTCAGAGTCTCGGACCAATTGCGCT 3587
Db 8526 CGGTGAGTCCGAAGCTACCGACGGAATGCGAAGATTCAAGTTTCGCAAAATTGGCA 8585
QY 3588 ACAATCAGGAGGACATGCGGATGTACCAAAATCCGTTACATGTCTTTCAGACGACGAC 3647
Db 8586 TAGATTAAGATGACATATCATCTACACGAAGATTAAGTA-----CGACACGGGC 8636
QY 3648 ATGACATCAAGAAAGACAGATAGGAAATAGCTATCAGCATCTGCACTCCGCTC 3707
Db 8637 ACGCATTTAGAAATGCCGTCCGTCATCTTTGAAGTAGCCACTCCGAGACATGTTTCG 8696
QY 3708 GTCTTGGCCACAAGGGTACTCTGTAGCTCATATGCTTCCAGGTGACAGTGAACG 3767
Db 8697 TCCATGCAACATGGAACATTTCATCTGCAAGGTGCCACCGGGTGAATTCCTGCAGG 8756
QY 3768 TCAATATCAG---AGCGAGCATCTGGAATTCATGACACCGTGGAGAAAAGATCAGGA 3824
Db 8757 TCTGATTCAGAGAACCGAAGACGGGTCCGTGCTGAGAAATCAATATCATATGACC 8816
QY 3825 GGAAGTTTGTGAGAGAGAGATCTTTGCCACCCGTTCATGAAAAGCTGTAAAGT 3884
Db 8817 CTCACCGGTGGTAGAGAAAATTTCAATTAGACCACTATGAAAAGAGATCCCTT 8876
QY 3885 GCCACGTTTACGACTTGAAGAGACGTCTGCCGGTACATTAACATGACAGGCGAG 3944
Db 8877 GCACACTTATCAACACACACGAGAACCGTGGAGAAATGACATGATATGCTGC 8936
QY 3945 GCCACACGCGTAAATCTCTATCTGAGAGAAAGGTACAGCGAAGTATCACTTAAACAC 4004
Db 8937 CAGATAGCCGAGACAGACGTGCTATCAACGACATCTGGCAATGTAAAGTCA---CAG 8993
QY 4005 CTCTGCGAAGAACGTCACTTACGATGTAAAGTGTGCGACTACAGACAGATATCTGA 4064
Db 8994 TCGGAGAAAGAAAGTGAATAACAATGCACTGTGAAACCGAAGCGTTGGCACTACTA 9053
QY 4065 GCACCGCAAGAAATGACCGGCTGCACTAAGCAAAACAGTGAATGCTTCAAGAGCG 4124
Db 9054 ATTGCAATGACATCAACGATCACTG---TCTAATAGACAGTGTCACTGACGTGACG 9110
QY 4125 ACCAAAGAAATGGGTCTTCACTCGCCGATCTTATTAGGACACAGACCACTCACTGC 4184
Db 9111 ACCATTAAGAAATGCACTTCACTCTTTCGTCGAGAGCCGAGAACCGGCTAGAA 9170
QY 4185 AAGTAATATGACATTCATTCGCTTACACCGACAGATGTGCGGTTCCGTTAGCTC 4244
Db 9171 AAGGCAAGTCAATATCCATTCCTGTTGCAACATCACTGACAGTTCAATGCGCG 9230
QY 4245 ACACGCTTACAGTACGAAAGTGTTCAAAGGATCACTCTTCACTGACATGAGAC 4304
Db 9231 GCGAACCAACCGTATCAACGCAAAAGAAAGATGACATGCACTTCAACCAATATATC 9290
QY 4305 CAACATGTGACACAGAAATTTGGGCTGCGAGACAGACGACAGCAAGATGATTA 4364
Db 9291 CCAAGCTCTTTTCTTACGCACTGGGTGAGACCCGAGATACAGAGAAATGGGTGA 9350
QY 4365 CAGGCTTACATCCAGAAATTTTCTGTGGGCGAGAGGGCTGAGTACGATGGGTA 4424
Db 9351 CAGGCGGTGAGACGACCATACCCGTACCAATGAGCGGATGAGATCACTGCGGAA 9410

QY 4425 ACCATGAACCATGACAGATCTGGGCCCCAGAGATGCGACAGAGGCAACCATGATGCG 4484
Db 9411 ACAACGACCCAGTGAAGCTTTGTTCTCACTACCACTGAAGGAAACCGCACGCTGGC 9470
QY 4485 CGATGAGATCATCATCACTATATCATCGGATCCAGTCTACACTGTATGTGCTGT 4544
Db 9471 CGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9530
QY 4545 GTGTGTGCTCTTCTGATCTCTGTGAGGACCTGATATGACAGCTTGCATGCGCAAG 4604
Db 9531 TCGGATGAGCTTACTGCGCTTATGATGATCTTTCGCTGTGATGATGATGATGATGAT 9590
QY 4605 CAGAAAGACCTGCTGACGCGCAATGACGCTTGCAGGACGAAAGGATCCACAGCAT 4664
Db 9591 CCGCAGTAAGTGTGATCCCTTATGCTTAAACACAGAGCTGACATTCCTGAGACG 9650
QY 4665 TAGCGTTTGTGCTGATTCGCGCAACACGCTGAAACATTTGAGAACTTTGAAAC 4724
Db 9651 TGGGATATCTGTGTGCGCCCGCGGCGACCGAGTATGTGAGACATATGCGCT 9710
QY 4725 ATCTGTGTTTAAACAAACCGTTTCTGCGACAGTGTGATTCCTTGGCAGCGC 4784
Db 9711 ACTGTGGGACCAAAACCAAGCGTTGTTCTGTTGAGTTGGCGCCCTGTTGCTGCA 9770
QY 4785 TTGTTATCTGTTCCGCTGC-----TTTCATGCTGATGCTTTTAT 4829
Db 9771 TCTCATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9830
QY 4830 TGGTTGACGCGTCTGCTGCGGAGAGTAGACGCTTTCGAAATGCGACCATGTGCGAA 4889
Db 9831 TAGTGTACTAGACGCTGGGGCAACCGCAGAGCTTTCGAAATTCGATATGCGA 9890
QY 4890 ATGTTCCGGGATTCCTGATTAAGCGTTGTGCAAGCGGAGGTACGCGCACTTAAC 4949
Db 9891 ACGTGTGGGTTTCCGTTATAGGCTCACTTAAGGCGCAGATATAGGCGCCCTCACTT 9950
QY 4950 TGGAGATCAAGGCTGCTCATTCGGAATTAACACTTCACTTAACAGAGATGAGTACCT 5009
Db 9951 TGCAGATGACAGTGTGTTGAAACACAGCTCGAACCAACCTTAAATTTGAAATACATACT 10010
QY 5010 GCAAAATTCACACAGTCAATCTTCAACCAAGTAAATGCTCGGCTCCCTGAGTGA 5069
Db 10011 GTGAGTACAAAGAGTGTCCCGTCCGTAAGTGAAGTCTCGGCGCTCAGAGTGT 10070
QY 5070 AAGCATCTTAAAGCGGATTAACATGCGCGCTTTTGGCGGTGTACCTTTATGT 5129
Db 10071 CCACTAAAGAAAGCTGACTACCAATGCAAGGTTTACACAGGCGGTATCCGTTCAATGT 10130
QY 5130 GGGGAGGCGCAATGCTTCTGAGACGTGAGAACACAACTGAGTGAAGCGGTAGCTG 5189
Db 10131 GGGAGGGGCAATTTGCTTCTGAGACTCAGAAAACGCACTCAGGAGGCTTACGTG 10190
QY 5190 AGTTGCTCCAGACTGCACTATGATCAACGAGTGCATTAAGTTTACACAGCTGCTC 5249
Db 10191 ATGATGCAAGTATGAGGATGATGAGCATCTGTTTCAAGAGCCATACAGCATGCG 10250
QY 5250 TGAAGTGGGCTGCGTATGATATGCGCAACACACCGGACACTGAGATAGCTTTGTCA 5309
Db 10251 TGAAGGCAAGTGAAGGTTATGACGCAACCTTAACAGAGTGTGAGATTTACCTGA 10310
QY 5310 ATGGCGTCAAGCGAGTTCTCAAGGAGCTGAAGGTCATAGCAGGCGGATATGACGCG 5369
Db 10311 ACGGAGACATGCGCTGACATAGGGGATCTCAATGTTCAATTCGCGCGCTGTATCGG 10370
QY 5370 CTTTTCACCTTTGACCATTAAGTGTCTCATCGAAAGGGGCTTGTATCACTACGACT 5429
Db 10371 CTTGACCCCGTTTGAACAACATATGTGTGTAACAAGAGCAAGTGTCAATCAGACT 10430
QY 5430 TCCCTGAGTATGAGCTATGAACACAGAGCGTTTGGCGATATTCAGCATCTCGCTTG 5489
Db 10431 TCCCGCGTATCGGATCTGGCAACAGGGCGTTCGCGCAATTCAAAGCAGAACTGTG 10490

QY 5490 ATGCTACAGACATAGTAGCCCGGCGACTGACATACGGCTGCTGAGACCTTCTGTCAAGACA 5549
DB 10491 AGAGTAAGACCTGCTAGCGGAACACGGCACTGAGCTGGACAGCCCTTACCCGGCATGG 10550
QY 5550 TCCAGTCCCTCCACACCCAGAGATATCAGGGTATGAAATGTGAAAGAACTAGAGAC 5609
DB 10551 TCCAGTACCCGTACACACAGACACCTTCAAGGGTTCAATATTTGGTTAAAGGAAAAAGGA 10610
QY 5610 GACCCCTGCAAGAAACAGACACCATTTGATGTAATTAAGTGAAGCTGCGAGCGT 5669
DB 10611 CAGCCCTAAATACGAGAGCTCTTTTGGCTGCCAAATCAAAAGAACTGTACAGGGCA 10670
QY 5670 CTAACTGTGCTTACGGGACATCCCTATCTCGATTGACATCCCTGATGACGCTTTTGA 5729
DB 10671 TGAATGTGGCGGTGGGAAACATCCCTGTCTCATGAAATTTGCTGACAGCGCTTTTACC 10730
QY 5730 GATCATCAGAAATCACCACATTTTGAAGTTAGCTGACAGTACAGATGCTGATTTAT 5789
DB 10731 GCATTGTGAGGCGCGGACCATATGACCTGACCTTGACAGTGGCTACTGTACGCACT 10790
QY 5790 CTGACACTTTGGTGTCTCTAATTAACAGTACAAAGCTGACAGGAGGACATTTGC 5849
DB 10791 CTTGGATTTTGGCGGCGCTCTTACACTGACGTAACAAAGCAAGGAGGAGCTGCT 10850
QY 5850 CAGTCTACTCCACTCCACAGAGCTGTTTGAAGAAAGCAACACATGTAAGTGGCG 5909
DB 10851 CTGTACTCTGCACTTAACGTAAGTACTCTACAGAGGCGCACCAAAAGTGAAGACAG 10910
QY 5910 TAGGCAGATTAACATCACTTTTACACATGAGCCCAAGCAAAATTTATAGTTTGC 5969
DB 10911 CAGGTAGGTGACCTTAACATCTTCCAGCGCAAGGCAATCACTTTTGGTGGTGGC 10970
QY 5970 TATGGGAGAGAGTCACTGACATGCTGAATGAACCAACCGGCGGACCAACATATG 6029
DB 10971 TATGACAGTCTAGGCGCACCTGTTAGCGTGTGAGGCCCGGAAAGCAACATAGTCC 11030
QY 6030 GAGAACCACTAAATGACACCAAGATTCAGAGCGGAGTTTCCAAACATTTGGAAT 6089
DB 11031 CATATGCGGCTAGCACAGTAAGTATGTTTCCAGACATGTCCGGGACCGCATATCAT 11090
QY 6090 GCGTGTCTGACATGTTGGGGAGCATATCCCTCATTTGTGTAGACTTATAGTTGG 6149
DB 11091 GGGTGCAGAAATATCGGGTGTCTGGGGGCTTGCATCGGCGCTATCCGTGGTGG 11150
QY 6150 T 6150
DB 11151 T 11151

RESULT 6
US-09-991-258-1
; Sequence 1, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanson, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991.258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902.537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216.995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1
; LENGTH: 12523
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-1
Query Match 14.3%; Score 926.2; DB 10; Length 12523;
Best Local Similarity 68.1%; Pred. No. 2.2e-265;
Matches 120; Conservative 0; Mismatches 613; Indels 6; Gaps 2;
QY 514 AACCAACAGAGTGTGCTAGAGCATCTGAGAGGCGGATTTGAGAGATCAATCCGTCAC 573
DB 5617 ACCCGCCAGGCGTAATTTGGGTATTTACAGAGAGAGATTTGAGGCGTTCTGACACAC 5676
QY 574 ACTCCACTGACGGTATGACCGGAGCGTATTTTCTCATTCGAAACAGGCCACGGTC 633
DB 5677 A--ACATGACGGTTGATGCGGGTGATCATCTTTTCTCCGACACCGGTCAAGGGC 5733
QY 634 ACCTTCAACAGAAATCAGTACGTCAATGTAACTCAAGAACTATATTGATCGGGCG 693
DB 5734 ATTTACAAACAAATCAGTAAGCAACGCTGTATCCGAGTGGTGTGAGAGAGACCG 5793
QY 694 TCCATGAGAGATTTAAGCGCGCGCGCTCGATCTCGAAGAGAGAAATGTACAGAGA 753
DB 5794 AATTGAGATTTGATGCTCCCGCGCTCGACCAAGAAAGAAATTAACACAGAGA 5853
QY 754 AACTCAATTTATGCGCTCTGAAAGAAATAGACAGGTATCAATCAGAAAGTAGAAA 813
DB 5854 AATTAGAGTTAAATCCACACCTGCTAACAAGAGATACAGATCCAGAAAGGTGAGA 5913
QY 814 ATATGAAGCAATTAACAGGAGGAGGAGTCTATTCTGATGGGCAATATCTATCATAG 873
DB 5914 ACATGAAGCCATTAACAGCTAGAGTATTCGCAAGGCGCTATGAGGATATTGAAAGCAG 5973
QY 874 AATGAATCCGTGAGGTATACAGAGTCAATTTCCGTGCAATCTACATCTGCAAGG 933
DB 5974 AAGGAA--AGTGAAGTCTACCGAACCTGCAATCTGTTCTTTGATTAATCTATG 6030
QY 934 TAATTAACAGGTTTACATCTGACAGAGTCCGGTTAAACGTGCACTTATTCGAAG 993
DB 6031 TGAACCGTCTTTTCAAGCCCAAGGTGAGAGGAGTGAAGCTGTAAAGCCATGTTGAAG 6090
QY 994 AGAATTAACCTACAGTACGACGATTTATGTATTAACAGATGAATACGATCGATCTTGACA 1053
DB 6091 AGAATTTCCAGATGTGGCTTCTTACTGTATTTCCAGAGTACGATGCTATTTGAGACA 6150
QY 1054 TGGTGAAGCGGCAATCGTCTGTCTATATACAGCACCTTTTTCGGCTAACTGAGAA 1113
DB 6151 TGGTTGAAGGAGCTTCAATGCTGTGACACTGTGCAATTTTTCCTGCAAGAGTGGCCCA 6210
QY 1114 GCTACCCAAAGACATATGATTTTGGAGCCAGAGATTAAGATACGCGGTCCATCGCCTTA 1173
DB 6211 GCTTTCCAAAGAAACATCTCTATTTTGAACCAACATACGATCGGAGTGTCTTCAAGCA 6270
QY 1174 TACAGAAATACATTAACAAATGTATTTGCTGACGTAATTAAGAAATTTCAACGTTACCC 1233
DB 6271 TCCAGAAACAGCTCCAGAACGTCCTGGAGCTGCAACAAAGAAATTTGCAATGTACAGC 6330
QY 1234 AAAATCCAGAAATTAACCTGTCTATATTTCCGGGCACTTAATGTGATTTTCAAGAAAT 1293
DB 6331 AAAATGAAGAAATTTCCGTATTTGATTTCCGGGCTTTTAAATGTGAATGCTTCAAGAAAT 6390
QY 1294 ACGATCAATGATGAGTACTGAGTACCTTTCGCGATTAACCTATTCGCTAATCTACAG 1353
DB 6391 ATGCTGTATTAATGATATTTGGGAAACGTTTAAAGAAACCCATCAGGCTTACTGAG 6450
QY 1354 AGAATGTTAGCAATATGTGCAAAAGCTGAAAGGCGGAAAGCAGACGATTTGTTGGA 1413
DB 6451 AAAACGTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6510

1414 ATACTATAAATCTAAACCGTTGAGAGATATACCAATGATCAATGCTATGATCTAA 1473
1415 AGACACATAATTTGAAATATGTTGACGAGCATACCAATGAGACAGGTTTGTATATGACTTAA 6570
1474 AGAGAGATGTCAAAATTACTCCCGCACGAAACATACAGAGAGAGCGGCTTAAGTGCAGG 1533
1475 AGAGAGACGTCGAAAGTACTCCAGGAAACAAACATCTAGAGAAACGCGCCCAAGGTACAGG 6630
1534 TTATTACGGGCTGAGATCCCGCTTCTCCGCTTACCTTTGGGGATCCATCGGGAATTAG 1593
1535 TGATCCAGGCTGCGGATCCGCTAGCAACAGGCTATCTGTGGGAATCCACCGAAGCTCG 6630
1594 TCCGTAGACTGAATGCGGTCTTCCCAATATCACTCTCTGACATGTACAGCGG 1653
1595 TTGAGAGATTAAATGCGGCTCTGCTTCCGAACATTTACATCTTTGATATGTGCGGTG 6750
1654 AAGATTTTATGATGATTTATGCTGATACATTTCCACACAGCGGACCCAGTATGGAAACGG 1713
1655 AAGACTTTGACGCTATTATAGCCGAGCACTTCCAGCCGCGGAGATTGTCTGGAAACTG 6810
1714 ACATCGGCTGTTGATTAAGGGAAGCAAGCGCTATCGCAATTCGCGCGTTGATGATCC 1773
1715 ACATCGGCTGTTGATTAAGGGAAGCAAGCGCTATCGCAATTCGCGCGTTGATGATTC 6870
1774 TTGAGACTTATGATGTCGACCAACCGCTCTTATGATGATGAGAGCGCGCTTCCGCAATA 1833
1775 TGGAAGACTTATGATGTCGACCAACCGCTCTTATGATGATGAGAGCGCGCTTCCGCAATA 6930
1834 TCACATCTGTCACCTTACTTACAGGAAACGAGCTTTAAATTTGGTCCATGATGAATCCG 1893
1835 TTTCATCAATCAATTTGCGCCACTTAAATTAATTTAAATTCGAGCCATGATGAATCTG 6990
1894 GTATGTTCTTAAGCGCTGTTGTCAACACACTAGTCAATATCATATGCTGAGAGTAC 1953
1895 GAATGTTCTCAACCTGTTGTGAACACATGATCAATTTATGATGCAAGAGTGT 7050
1954 TACGTGAACGCTTAAACAGCAGTCAAGCGGCTCTTATGAGGCAAGTAACTAGTGC 2013
1955 TGAGAGAAACGCTTAAACAGCAGTCAAGCGGCTCTTATGAGGCAAGTAACTAGTGC 7110
2014 ATGCTGCTGCTCCGACACCTTATGAGGCAAGTCACTTGGCTGAACATGGAAG 2073
2015 AAGAGTCAATTCGACAAATTAATGAGCAAGTGCAGCCTGCTGTAATATGGAAG 7170
2074 TAAATTTATGATGAGTATGATGATGATCAAAAGCACTTATGAGGAGGATTTATCC 2133
2075 TCAAGATTATGATGCTGCTGAGGCAAGGCTCTTATGAGGAGGATTTATTT 7230
2134 TGTGTGACCAAGATTAACAGGCAAGCTTGCAGAGTCCGACCTCTTAAAGGCTTTTAA 2193
2135 TGTGTGACCTCGGTGACCGGCAAGGCTGCTGAGGCAAGCCTCTTAAAGGCTTTTAA 7230
2194 AGCTTTGAAAACATTTGCAAGTCAATGATCAAGAGTGCAGCGCGCGGCACTGC 2253
2195 AGCTTTGAAAACCTCTGCGACAGACGATGAAATGATGACAGAGGAGGCACTTGC 7350
2254 ATGATGAAGCAATGAGTGAAGCAAGTGAATTTAGCGAGAGTGAAGGCGGTAG 2313
2255 ATGAAAGTCAACACGCTGGAACCGAGTGGTATTTCTTCAAGAGTGTGCAAGGAGTAG 7410
2314 AATCAGATATGAGATCATCTGAGGCTGATCATCATCTGTCACAGTTAGCCG 2373
2315 AATCAGATATGAGATCATCTGAGGCTGATCATCATCTGTCACAGTTAGCCG 7470
2374 AAGCGTTAAGAACTTCAAGAGCATTAAGAGGAGCCCAATCACTCTTACGCTGACCTA 2433
2375 GCACTGTTAATTAATCAATCACTGAGTACCTGAGAGGCGCCCTATATCTCTACGCTAACCTG 7530
2434 AATAGTGAAGCTTATGAGC 2452
7531 AATGAGCTTACGACATAGTC 7549

RESULT 7
US-09-991-258-14
Sequence 14: Application US/09991258
Patent No. US20020141975A1
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Maughan, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swanson, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: 01113.000103
CURRENT APPLICATION NUMBER: US/09/991,258
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902,537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,995
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 12379
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-14
Query Match 14.3%; Score 924.6; DB 10; Length 12379;
Best Local Similarity 68.0%; Pred. No. 6.5e-265; Indels 6; Gaps 2;
Matches 1319; Conservative 0; Mismatches 614;
514 AACCAACAGAGTCTAGAGAGCATCTCAGAGAGGAGTTGACGAGTACATCCGTCAAC 573
517 ACCCGCAGCGTAAATAGGATGATTAAGAGAGAGGTTGAGCGCTTGTAGACACAC 5676
574 ACTCAACTAGCGGTATGAGAGGAGCGTATTTCTCATGCGAAACAGGCCAGTGC 633
577 A---ACAATACGCTTGTGATGCGGTGCATATACATCTTCTCCGACACCGGTCAGAGGC 5733
634 ACCTTCAACGAATTCAGTGTGATGATTAACATCAAGAACATATTTGGATGGGCGG 693
635 ATTACAAACAAATTCAGTGAAGGCAACGCTGCTATCCGAAAGTGTGAGAGGACCG 5793
694 TCCATGAGAGTATTTAGCGCCCGGCTCGATCTCGAAGAGAGAAATGTTACAGAGA 753
695 AATTGAGATTTCTGATGCGCCCGGCTCGACCAAGAAAGAAATTTACTACGCAAGA 5853
754 AACTGAATTAATGCGCTCTGAAAGAAATGAAAGCAGATATCAATCAAGAAAGTAGAAA 813
755 AATTACAGTTAAATCCACACCTGCTAAGAGCAATGATCCAGAGAGGTGAGA 5913
814 ATATGAAGCAATTAACGCGGAGGAGTCAATTTCTGATGATGGGACATATCTATCATCAG 873
815 ACATGAAGGCAATTAACGCTGATGATTTCTGCAAGGCTGAGGCAATTTTGAAGCGAG 5973
874 AAGTGAATCTGTCGAGTGTATTAAGAGTCAATTTCTGTAACCAATCTACTGTCAGCG 933
875 AAGGAGTCTGTCGAGTGTATTAAGAGTCAATTTCTGTAACCAATCTACTGTCAGCG 993
934 TAAATTAACAGTTTATCATCTGACAGAGTGCAGGTTAAACGTCACATTAATTTCAAG 6030
935 TGAACGTCCTTTTCAAGCCCAAGGTCGACAGTGAAGGCTGTAAGCCCATGTTGAAG 6090
994 AGAATTTACCTACAGTACGAGTATTTGATTAACAGATGAATGAGATGCGGATTTGACA 1053
995 AGAATTTCCGACGTGCTTTCTTACTGATTAATTTCAAGATGAGTACGATGCTTATTTGACA 6150

1054 TGTGTGACGCGCATCGTGTCTAGATACAGCCATTTTGTCCGGCTAACTGAGAA 1113
1113 TGTGTGACGCGCATCGTGTCTAGATACAGCCATTTTGTCCGGCTAACTGAGAA 1113
6151 TGTGTGACGCGCATCGTGTCTAGATACAGCCATTTTGTCCGGCTAACTGAGAA 6210
1114 GCTACCCAAAGAGATAGCTATTGTCAGCCAGAGATTAAGATCAGCCGCTCCATGCGCTA 1173
6211 GCTTCCAAAGAGATAGCTATTGTCAGCCAGAGATTAAGATCAGCCGCTCCATGCGCTA 6270
1174 TACAGATACATTAACAAATGTAATGGCTGACGCTACTAAAGAAATGCAACGTTACCC 1233
6271 TCCAGAACACGCTCCAGAACGCTGTCGAGCTGCCACAAAGAAATGCAATGTCACGC 6330
1234 AAATGCGAAATTAACCTGCTTGAATGTCGCGGCAATTAATGTTGATTTTCAAGAAAT 1293
6331 AAATGCGAAATTAACCTGCTTGAATGTCGCGGCAATTAATGTTGATTTTCAAGAAAT 6390
1294 ACGATGCAATGATAGTACTGAGATACCTTTGCGGATTAACCTTATTGGGCTAACTACAG 1353
6391 ATGCGTGAATTAATTAATTAATGGAACGTTTAAAGAAACCCCATCAGGCTTACTGAAAG 6450
1354 AGAAGTACGCAATATGATGACAAAGCTGAAGAGGCGGAAAGCAGACGATTTGTCGA 1413
6451 AAAAGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6510
1414 ATACTCATTAATCTAAACCGTTGACAGAGATACCAATGATCAATTCGTCATGATCTAA 1473
6511 AGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6570
1474 AGAGAGATGTCAAAGTACTCTCCGCGACGAAACATTAACAGAGAGCGGCTTAAGTGCAG 1533
6571 AGAGAGATGTCAAAGTACTCTCCGCGACGAAACATTAACAGAGAGCGGCTTAAGTGCAG 6630
1534 TTATTCAGGCTGACAGATCCCGCTGCTACCGCTTACCTTGGGATTCATCGGGAATAG 1593
6631 TGATTCAGGCTGACAGATCCCGCTGCTACCGCTTACCTTGGGATTCATCGGGAATAG 6690
1594 TCCGTAAGTGAATGCGGTGCTTTCGCCAAATATCATCTCTTCCACATGTCAGCG 1653
6691 TTAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6750
1654 AAGATTTGATGCGATTAATGCTGAACATTTTCACACGCGCGACCCAGTATGGAACCG 1713
6751 AAGATTTGATGCGATTAATGCTGAACATTTTCACACGCGCGACCCAGTATGGAACCG 6810
1714 ACATGCGCTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1773
6811 ACATGCGCTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6870
1774 TTGAGAGCTAGAGTGCAGCAACGCTCTTGAATTTGATAGAGCGGCTTCCGCAATA 1833
6871 TGGAGAGCTAGAGTGCAGCAACGCTCTTGAATTTGATAGAGCGGCTTCCGCAATA 6930
1834 TCAATCTGTGCACTTACCTACAGAACGAGTTTAAATTTGTCGCAATGATGAATCCG 1893
6931 TTTCAATCAATCAATTTGCCCCACTAAATTAATTAATTTGCGAGCCATGATGAATCCG 6990
1894 GTATCTTTTAAACGCTTTTGTGCAACACACTAGTCAATATCAATGTTCTGACAGAGTAC 1953
6991 GAATTTTCTCAACACTGTTTGTGCAACACACTAGTCAATATGATTTCCACAGAGAGTGT 7050
1954 TACGTAACGAGTGAACAGTCAAGGCTGCGGCGCTTATCGGCGACAGTAACATATGTC 2013
7051 TGAAGAGAGGCTTAACCGGATCAATGTGACGACTTCAATGAGATACATATGCTGA 7110
2014 ATGCTGTGCTCCGACACTTGTATGCGAGAGATGCGCACTTGCTGAACATGAGAG 2073
7111 AAGGAGTCAAAATCGCAAAATTAATGCGAGAGAGTGGCGCACTGGTTGAATGAGAG 7170
2074 TAAATTAATTAATGATGATTAATGATGATCAATCAATCAATCAATCAATCAATCAATCA 2133
7171 TCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7230
2134 TGTGTGACGCGCATCGTGTCTAGATACAGCCATTTTGTCCGGCTAACTGAGAA 2193

1113 TGTGTGACGCGCATCGTGTCTAGATACAGCCATTTTGTCCGGCTAACTGAGAA 1113
1113 TGTGTGACGCGCATCGTGTCTAGATACAGCCATTTTGTCCGGCTAACTGAGAA 1113
6210 TGTGTGACGCGCATCGTGTCTAGATACAGCCATTTTGTCCGGCTAACTGAGAA 6210
1173 GCTACCCAAAGAGATAGCTATTGTCAGCCAGAGATTAAGATCAGCCGCTCCATGCGCTA 1173
6270 GCTTCCAAAGAGATAGCTATTGTCAGCCAGAGATTAAGATCAGCCGCTCCATGCGCTA 6270
1233 TACAGATACATTAACAAATGTAATGGCTGACGCTACTAAAGAAATGCAACGTTACCC 1233
6330 TCCAGAACACGCTCCAGAACGCTGTCGAGCTGCCACAAAGAAATGCAATGTCACGC 6330
1293 AAATGCGAAATTAACCTGCTTGAATGTCGCGGCAATTAATGTTGATTTTCAAGAAAT 1293
6390 AAATGCGAAATTAACCTGCTTGAATGTCGCGGCAATTAATGTTGATTTTCAAGAAAT 6390
1353 ACGATGCAATGATAGTACTGAGATACCTTTGCGGATTAACCTTATTGGGCTAACTACAG 1353
6450 ATGCGTGAATTAATTAATTAATGGAACGTTTAAAGAAACCCCATCAGGCTTACTGAAAG 6450
1413 AGAAGTACGCAATATGATGACAAAGCTGAAGAGGCGGAAAGCAGACGATTTGTCGA 1413
6510 AAAAGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6510
1473 ATACTCATTAATCTAAACCGTTGACAGAGATACCAATGATCAATTCGTCATGATCTAA 1473
6570 AGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6570
1533 AGAGAGATGTCAAAGTACTCTCCGCGACGAAACATTAACAGAGAGCGGCTTAAGTGCAG 1533
6630 AGAGAGATGTCAAAGTACTCTCCGCGACGAAACATTAACAGAGAGCGGCTTAAGTGCAG 6630
1593 TTATTCAGGCTGACAGATCCCGCTGCTACCGCTTACCTTGGGATTCATCGGGAATAG 1593
6690 TGATTCAGGCTGACAGATCCCGCTGCTACCGCTTACCTTGGGATTCATCGGGAATAG 6690
1653 TCCGTAAGTGAATGCGGTGCTTTCGCCAAATATCATCTCTTCCACATGTCAGCG 1653
6750 TTAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6750
1713 AAGATTTGATGCGATTAATGCTGAACATTTTCACACGCGCGACCCAGTATGGAACCG 1713
6810 AAGATTTGATGCGATTAATGCTGAACATTTTCACACGCGCGACCCAGTATGGAACCG 6810
1773 ACATGCGCTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1773
6870 ACATGCGCTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6870
1833 TTGAGAGCTAGAGTGCAGCAACGCTCTTGAATTTGATAGAGCGGCTTCCGCAATA 1833
6930 TGGAGAGCTAGAGTGCAGCAACGCTCTTGAATTTGATAGAGCGGCTTCCGCAATA 6930
1893 TCAATCTGTGCACTTACCTACAGAACGAGTTTAAATTTGTCGCAATGATGAATCCG 1893
6990 TTTCAATCAATCAATTTGCCCCACTAAATTAATTAATTTGCGAGCCATGATGAATCCG 6990
1953 GTATCTTTTAAACGCTTTTGTGCAACACACTAGTCAATATCAATGTTCTGACAGAGTAC 1953
7050 GAATTTTCTCAACACTGTTTGTGCAACACACTAGTCAATATGATTTCCACAGAGAGTGT 7050
2013 TACGTAACGAGTGAACAGTCAAGGCTGCGGCGCTTATCGGCGACAGTAACATATGTC 2013
7110 TGAAGAGAGGCTTAACCGGATCAATGTGACGACTTCAATGAGATACATATGCTGA 7110
2073 ATGCTGTGCTCCGACACTTGTATGCGAGAGATGCGCACTTGCTGAACATGAGAG 2073
7170 AAGGAGTCAAAATCGCAAAATTAATGCGAGAGAGTGGCGCACTGGTTGAATGAGAG 7170
2133 TAAATTAATTAATGATGATTAATGATGATCAATCAATCAATCAATCAATCAATCAATCA 2133
7230 TCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7230
2193 TGTGTGACGCGCATCGTGTCTAGATACAGCCATTTTGTCCGGCTAACTGAGAA 2193

RESULT 8
US-09-991-258-17
Sequence 17, Application US/09991258
Patent No. US20020141975A1
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Caley, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Swanson, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: 01113.000103
CURRENT APPLICATION NUMBER: US/09/991,258
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902,537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,995
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO. 17
LENGTH: 13584
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-17

Query Match 14.3%; Score 924.6; DB 10; Length 13584;
Best Local Similarity 68.0%; Pred. No. 7e-265;
Matches 1319; Conservative 0; Mismatches 614; Indels 6; Gaps 2;
514 AACCAACGAGAGTGGCTAGACGATCTGAGAACGCGAATTTGAGACGAGTACATCCGTCAAC 573
5617 ACCGCGCAGGCTAAATGAGGTGATTAACAGAGAGAGTGGAGCGTTGCTGACACAAAC 5676
574 ACTCAACTGACGATGATGAGAGGAGGCTATTTTCTCATCGGAACAGGCCAAGGTC 633
5677 A--ACATGACGCTTTGATGAGGAGGCTATCATCTTTCTCCGACACCGGTCAAGGCG 5733
634 ACCTTACAGAGAAATGAGTACGTAATGTAATTAATCAAGAAACCTATATTTGATGAGGCGG 693
5734 ATTACAAACAAATTAAGTAAAGCAACGCGTATTCGAAAGTGTGTTGAGAGAGACCG 5793

694 TCCATGAGATATTAGCCCGCGCTCGATCTCGAAAGAGAGAAATGTACAGAGA 753
5794 AATTGGAGATTTCTATGCCCCCGGCTCGACCAAGAAAGAAATTTACTAGCAAGA 5853
754 AACTGCAATTATGCGCTCTGAAAGAAATAGACAGGATCAATCAAGAAATAGAAA 813
5854 AATTACAGTTAAATCCACACCTCTAAACAGAACAGATACCACTCCAGGAAGGTGAGA 5913
814 AATTGAAGCAATTACGCGGAGGAGCACTATTCTGATTTGGGACATCTATCATCAG 873
5914 ACATGAAGGCTATACAGCTAGACGTATCTGCAAGGCTAGGGCATTTATTTGAAGCG 5973
874 AAGTAATCTCTGAGATGTATACAGATCAATTAATCTGTACCAATCTACTCTCAAG 933
5974 AAGGAA--AGTGAATGTCTACAGAACCTGCACTCTGTTCTTTGATTTCACTAGTG 6030
934 TAATTAACAGTTTACATCTGCAAGGTGCGGTTAAACGTGCACTTATGTTACAG 993
6031 TGAACCGTCTTTCAAGCCCAAGGTGCGAGTGAAGCCTGTAAAGCCATGTTGAAG 6090
994 AGAATTACCTAGTAGTCCAGTATGTATACAGATGAATAGATGCGTATCTTACA 1053
6091 AGAATTTCCGACTGTGCTCTTACTGTATTTCCAGATGAGATGCTATTTGAGCA 6150
1054 TGGTGAAGCGGCGCATCTGCTGTCTAGATACAGCCACTTTTGTCCGGCTAACTGAGA 1113
6151 TGGTGAAGGAGCTTCACTGTCTGTAGACACTGCACTGATTTTGGCCCTGCAAGAGTGGCA 6210
1114 GCTAACCAAGAGCATAGCTATTTGCAAGGATGATGATGACCGGCTCCATGCGCTTA 1173
6211 GCTTTCAAGAAACACTCTCTATTTGGAACCAATACATCCGAGTGTCTTCAACGA 6270
1174 TAGAGATACCTTCAAAATGTATTTGCTGAGCTGACTTAAAGGAATTCGAACCTTACC 1233
6271 TCCGAACACGCTCCAGAACTCTGAGCTGCACAAAAGAAATGCAATGTCAGC 6330
1234 AATGCGAGAAATTAAGTCTTAAATTCGCGGCACTTAAATGTTGTTTCAAGAAAT 1293
6331 AATGAGAGAAATTCGCGGATTTGGAATTCGCGGCTTAAATGTTGGAATGTTCAAGAAAT 6390
1294 ACGCATCAATGATGAGTACTGAGATACCTTTGCGCATTAACCTTATTCGGCTAACTAG 1353
6391 ATGGGTAAATTAAGATATTTGGGAAAGCTTTAAAGAAACCCCATGAGCTTACTAGAG 6450
1354 AGAAGCTTACGCAATATGTGCAAGCTGAAGGCGGAAAGGAGAGCAAGCATTTGTTGCA 1413
6451 AAAACGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6510
1414 ATAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1473
6511 AGACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6570
1474 AGAGAGATGTCAAAGTTACTCCGCGCAAGAAACATACAGAGAGCGGCTTAAAGGTGAGG 1533
6571 AGAGAGAGCTGAAGGATCTCCAGAAACAAACATACAGAGAGCGGCTTAAAGGTGAGG 6630
1534 TTATTCAGGCTGAGATCCCTTGTACCTGCTTACCTTTGGGGAATTCATTCGGGAATTAAG 1593
6631 TGATTCAGGCTGCGGATCCGTAGCAACAGGATCTGTGCGGAATTCACCGAAGAGCTGG 6690
1594 TCCGTAAGTGAATGCGGCTTGTGCAAAATATCACTCTCTTGCACATGTCAGCGG 1653
6691 TTAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6750
1654 AAGATTTGATGCGATTAATGCTGAACATTTCCACAGCGGAGCCAGATTTGGAACCGG 1713
6751 AAGACTTTGACGCTATTAATAGCCAGACCTTCCAGCTCGGGGATTTGTTGTTCTGGAACCTG 6810
1714 ACATGCGGCTGTTGATTAAGGAGCAAGCAAGCATGCTATTCGCGCTTGAATGATCC 1773
6811 ACATGCGGCTGTTGATTAAGGAGCAAGCAAGCATGCTATTCGCGCTTGAATGATTC 6870
1774 TTGAGACTTATGATGTCGACCAACCGCTCTTAAGATTTGATAGAGCGGCGTTGCGCATATA 1833

6871 TGGAGACTTATGAGTGTGAGCAGAGCTGTGACGCTGATTTGAGCGGCTTTGCGCGAAA 6930
1834 TCACATCTGTCGACCTACCTACAGAGACGAGTTTAAATTTGGTCCATGATGAAATCCG 1893
6931 TTTCACTCAATTAATTTGCGCCACTTAAATTTAAATTTGAGAGCCATATGAAATCTG 6990
1894 GTATGTTCTTAAGGCTGTTTGTCAACACACTAGTCAATATCATGATTTGTAGCAGAGTAC 1953
6991 GAATGTTCTCAACACTGTTTGTGAACAGACTCATTAATGATTTGATGCAAGAGAGTGT 7050
1954 TACGTAAACGTTAACCACGTCAGCGGTGCGGCTCTATGCGCAGATPAACATAGTGC 2013
7051 TGAGAGAACGCTTACCGGATCAACATGTCACACATTCATTTGAGATGACAAATATCGTGA 7110
2014 ATGCTGCTCTCCGACACTTATGCGGAGAGATGCGGCACTGCTGTAACATGGAAG 2073
7111 AAGAGTCAATGCGAATAATTAATGAGCAGACAGTGGCCACCTGTTAATATGGAAG 7170
2074 TAAATATTTATGATGAGTATTTGATTAATCAAAACACCTTCTGTTGGGAGATTATCC 2133
7171 TCAGATTTATGATGCTGTGTGTGGCGAAGAACGCTTATTTCTGTGAGGTTATTT 7230
2134 TGGTGAACGATPAACAGGACAGCTGACAGATGCGAGACCTTAAAGGCTTTTA 2193
7231 TGTGTACTCCTGACCGGACAGCGTGCCTGTGGGAGACCCCTTAAAGGCTGTTTA 7290
2194 AGCTTGAAGAACATTTGCGGATGATGATATCCAAAGCTGCAAGCGGCGGCACTGC 2253
7291 AGCTTGGCAACCTCTGCGAGCAGACGATGAAATGATGATGATGATGATGATGATGATG 7350
2254 ATGATGAAGCAATGCGATGAGAGAGAAATTTGAATTAACGAGATTAAGTAAAGGCTGAG 2313
7351 ATGAGAGTCAACAGCTGAGAACGATGAGTATTTCTTAAGAGCTGTGAAGGAGTAC 7410
2314 AATTCAGATPAACGATCATATCTGCGAGGCTGTATCATGCTCTGTGCAAGTTTAA 2373
7411 AATCAAGTATGAAACCGTGAAGACTTCCATCATGATTTAGGCAATGATCACTAGCTA 7470
2374 AAGCGTTAAGAACTTCAAGAGCATTAAGAGGAGCCCAATCACCTCTTAAAGGCTGACCTA 2433
7471 GCAAGTGAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7530
2434 AATAGTGAAGCTAGTAGAC 2452
7531 AATGACTACGACATAGTC 7549

RESULT 9
US-09-991-258-2
Sequence 2, Application US/09991258
Patent No. US20020141975A1
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryden, Sergey
APPLICANT: Cailey, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swanson, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: 01113.000103
CURRENT APPLICATION NUMBER: US/09/991.258
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902.537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216.995
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2

LENGTH: 7479
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
NAME/KEY: CDS
LOCATION: (1)...(7479)
US-09-991-258-2

Query Match 14.1%; Score 916.2; DB 10; Length 7479;
Beet Local Similarity 68.2%; Pred. No. 1.5e-262;
Matches 1304; Conservative 0; Mismatches 603; Indels 6; Gaps 2;

QY 514 AACCAACGAGATCGCTAGAGGAGATCTCAGAACGGAGTTGAGCGAGTACATCCGTCAC 573
DB 5573 ACCCGCCAGCGTAAATAGGGGATTTACAGAGAGAGATTGAGGCTTCTGATGACACAC 5632
QY 574 ACTCCAACTGACGGTATGAGAGCGGAGCGGTATTTCTCATCGAAGACGCGCAAGTTC 633
DB 5633 A---ACAATGACGGTTTGAATGGGGGATACATCTTTTCTCCGACACGGGTCAAGGCG 5689
QY 634 ACCTTCAACAGAAATCAGTACGTCAATGAACTACAGAACCTATATTGATCGGGCCG 693
DB 5690 ATTTCACACAAAATCAGTAAAGCAACGAGTCTATCCGAGTGGTGTGAGAGGACCG 5749
QY 694 TCCATGAGAGTATTACGCGCGCGCTGATCTGAAAGAGAGAAATGTTACAGAAAG 753
DB 5750 AATTGAGATTTTCTATGCGCCGCGCTGACCAAGAAAGAGAAATTAATACGCAAGA 5809
QY 754 AACTGCAATTAATCGCTCTGAGAGAAATAGAGAGAGTATCAATCACGAAAGTAGAA 813
DB 5810 AATTACAGTAAATCCCAACCTGCTAACAGAGAGATACCAAGTCCAGAGAGTGGAGA 5869
QY 814 ATATGAAAGCAATTAACAGGAGCGAGCTATTTCTGATGGGACATATCTATCATAG 873
DB 5870 ACATGAAAGCCATTAACAGTACAGTATTTCTGCAAGGCGCTAGGGCATTAATTTGAAGGCG 5929
QY 874 AAGTAACTCTGTGAGTGTACAGAGTCAATTAATCTGTACCAATCTATCTGCAAGG 933
DB 5930 AAGGAAA---AGTGAAGTCTACCGAACCTGTGATCTGTTCTTTGTAATTAATCTAGTG 5986
QY 934 TAATTAACAGGTTTACATCTGCAAGAGTCCGCTTAAACCTGCAACTAGTTATCAAG 993
DB 5987 TGAACCGTGCCTTTTCAAGCCCAAGGTGCGAGTGAAGCCGTAAACGCAATGTTGAAG 6046
QY 994 AGAATTAACCTTACAGTACGCAATTAATGTAATACAGATGATACATGCGTATCTTGACA 1053
DB 6047 AGAATCTTCCGACTGTGGCTTCTTACTGATTAATTCAGAGATGATGCTATTTGGACA 6106
QY 1054 TGGTGAAGGCGCATCGTGTCTGTAGATACAGGCACTTTTGTCCGCTAACTGAGAA 1113
DB 6107 TGGTGAAGGCGCATCGTGTCTGTAGATACAGGCACTTTTGTCCGCTAAAGCTGCGCA 6166
QY 1114 GCTTACCAAGAGATAGCTATTGTCAGCCAGAGATAGATACAGCCGCTCCATCGCTTA 1173
DB 6167 GCTTCCAAAGAAACACTCTATTGTCAGCCAGATACGATGCGGAGTCCCTTAAGCA 6226
QY 1174 TACAGATACATTAACAAATGTAATGTCGTCAGCTACTAAAGAAATGCAAGTTACCC 1233
DB 6227 TCCAAAGACGCTCAGAACGTCCTGCGAGCTGCACAAAAGAAATGCAATGTCAGCG 6286
QY 1234 AAATGCGAAGATTAATCTGTCTTAAGTGGCGGCACTTAATGTTATGTTTCAAGAAAT 1293
DB 6287 AAATGAGAAATGCGCTGATGATGATGCGGCGCTTAATGTTGAAGTCTTCAAGAAAT 6346
QY 1294 ACGCATGCAATGATAGTACGAGTACCTTTGCGGATACCTTATGCGGCTACTAGAG 1353
DB 6347 ATGCGTGAATTAATTAATGTTGAAGCGTTAAAGAAACCCCAATGAGGCTTACTGAG 6406
QY 1354 AGAAGCTTAACGCAATATGTGACAAAGCTGAAGGCGGCAAGACAGAGCATTTGTTGGA 1413
DB 6407 AAAAGGTGTAATTAATTAATTAATTAAGAGCAAAAGCTGCTGCTTTTGGGA 6466

QY 1414 ATACTAATATCTAAACCGTTGACAGAGATACCAATGATCAATTCGTATGATCTTA 1473
DB 6467 AGACATATATTTGAATATGTTGACGACATACCAATGACAGGTTTGAATGACTTAA 6526
QY 1474 AGAGAGATGCAAAATTAATCTCCCGACGAAACATACAGAGAGCGGCTTAAGTTCAGG 1533
DB 6527 AGAGAGAGTGAAGATGACTCCAGAGAACAAACATATCTGAGAGAGCGGCCCAAGTACAG 6586
QY 1534 TTATTCAGGCTGCAATCCCTTGTCTTACCGCTTACTCTTGGCGGATCCATCGGGAATTAG 1593
DB 6587 TGATCCAGGCTGCGCATCCGCTTACGCAACAGGCTATCTGTGCGGAATCCACCGAGCTGG 6646
QY 1594 TCCGTAAGCTGAATCGGCTGCTTGCACAAATATCAATCTCTTCCGATGTCAGCGG 1653
DB 6647 TTAGAGATTAATTCGGCTCTGCTTCCGAACTTATCACTGTTTGAATATGCGGCTG 6706
QY 1654 AAGATTTTGAATGCAATTAATGCTGAACATTTCCACACGCGCACCCAGTATTGAAACGG 1713
DB 6707 AAGACTTTGACGCTATTAATGACGAGCACTTCCAGCTGGGGAATTGTCTGGAACCTG 6766
QY 1714 ACATGCGTCTGTTGATTAATTAAGCGAAGACGCTATGCGCATTTGCGGCTTGAATTC 1773
DB 6767 ACATGCGTCTGTTGATTAATTAAGTGAAGACGACGCTATGCTGACCGGTTAATGATTC 6826
QY 1774 TTGAGACTTGAATGTCGACCAACCGCTTATGATTTGAGAGGCGGCTTGGCAATA 1833
DB 6827 TGGAGACTTGAATGTCGACCAACCGCTTATGATTTGAGAGGCGGCTTGGCAATA 6886
QY 1834 TCACATCTGTGACCTTACCTACAGAGACGAGGTTTAAATTTGCTGCAATGATGAATCG 1893
DB 6887 TTTTATCAATATCATTTGCGCACTAAATTAATTAATTTGAGACCAATGATGAATCTG 6946
QY 1894 GTATGTTCTTAAAGCTGTTTGTCAACACATAGTCAATATATATGTTGTACAGAGTAC 1953
DB 6947 GAATGTTCTTCAACATGTTTGTGAACACAGTATTAATTAATGTTATTCACAGCAAGTGT 7006
QY 1954 TACGTAACGCTTAAACAGCTCAGGCTGCGGCGCTTATGCGGACGATTAATGATGTC 2013
DB 7007 TGAAGAAACGCTTAAACGAGATCAATATGACATCTTATGAGATGACATATCTGTA 7066
QY 2014 ATGCTGTCTCCGACACTTGTATGAGGAGAGATGCGCACTTGGCTGAACATGAG 2073
DB 7067 AAGGATCAATTCGACAAATTAATGAGACAGAGTGGCCACTGTTGAATATGGAAG 7126
QY 2074 TAAATTAATTAATGATGATTAATGTTATCAAAAGCACTTCTTGTGGGGATTTATCC 2133
DB 7127 TCAAGATTAATGATGCTGTGTGGGCGAGAAAGCCCTTATTTCTGTGAGGCTTATTT 7186
QY 2134 TGGTGAACGATTAACAGGCAAGCTGACAGTGCAGAGCTCTTAATAAGGCTTTTAA 2193
DB 7187 TGTGTGACTCCGTAACCGGCAAGGCTGCGGCTGAGAGAGCCCTTAATAAGGCTTTTA 7246
QY 2194 AGCTTGAAGAAACATATGCGAGTGTATGATACCAAGCTGCGACCGCGGCACTGTC 2253
DB 7247 AGCTTGAAGAAACCTTGGCAGACAGATGAAATGATGATGACAGAGAGAGGCAATGTC 7306
QY 2254 ATGATGAACATGCGATGAGAAACAGATTTGAATTAACGAGAGTGAAGTGAAGGCGTATG 2313
DB 7307 ATGAAGAGTCAACAGCTGGAACCGAGTGTATTTCTTCAAGAGCTGTCAGAGGCAATG 7366
QY 2314 AATCAGATGAGAGATCAATCTGAGGCGCTGATCATACGCTCTGTCACGTTACCG 2373
DB 7367 AATCAGAGTGAAGACCGTGAAGAAATTCATCATATGTTATGCGCATGACTACTTAGCTTA 7426
QY 2374 AAAGCTTAAGAACTTCAAGAGCAATTAAGAGGAGCGCCATCACTCTTAAGCGC 2426
DB 7427 GCAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7479

RESULT 10
US-09-507-362-102
; Sequence 102, Application US/09507362

Publication No. US20030096397A1
GENERAL INFORMATION:
APPLICANT: Dubensky Jr., Thomas W.
POL: John M.
Bell, Barbara A.
Schlesinger, Sondra
Dryga, Sergey A.
Frolov, Ilya
TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
SYNTHESIS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group PLLC
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,362
FILING DATE: 18-Feb-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 930049.457D6 /1196.011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 8000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-507-362-102
Query Match 12.8%; Score 827.6; DB 11; Length 8000;
Best Local Similarity 61.9%; Pred. No. 5.9e-236;
Matches 1390; Conservative 0; Mismatches 829; Indels 25; Gaps 4;
QY 580 ACTGACGGTATGAAGCGGAGCGTATATTTCTCATCGAAGAGCCCAAGTCACCTTC 639
DB 5746 ACTGACTCAACCGGGGTAGTGGTACATATTTTCGACGACACAGGCCCTGGGACATTGC 5805
QY 640 AACGAATCACTACGTCAATGTAACTACAGAACCTATATTGGATCGGCCCTGCATG 699
DB 5806 AAAAGAGTCGGTCTGACAGAACGCTTACAGAACGACCTTGAGGCGCATATGCTCGG 5865
QY 700 AGAAGTATACGCCCCCGCTCGATCTCGAAAGAGAAATGTATACAGAAAGAACTGC 759
DB 5866 AAAAATTCATGCCCGGTGCTGACAGCTGCAAGAAAGAACTCACTCACTCGTACC 5925
QY 760 AATTATGCGCTCTGAAGAAATAGAACAGTATCATACGAAAGTAGAATAATAGA 819
DB 5926 AGATGATGCCCAACCGAAGCCAAAGTAGTACCACTCTGTAAGTAGAATAATCAGA 5985
QY 820 AAGCAATTAACGAGAGGAGCTATTTCTGATTTGGGACATATCTATCATCAGAAAGTGA 879
DB 5986 AAGCATTAACCACTGAGGAGTACTGTACAGACTGATATACCTGTCACAGATC 6045
QY 880 ATCTGTGAGTGTATACAGATCAATTAATCTGTACCAATATATCTGCAACGGTAATTA 939
DB 6046 AGCCA--GAATGCTATTAAGATCACTATCGAAACCATTTGTACTCCAGTAGCGCCG 6102
QY 940 ACAGGTTTACATCTGACAGAGTCCGGTTAAACGTGCACTTATGTTATCCAGAGATT 999

DB 6103 CGAACTCTCGATCCACAGCTTGCTGTAGCTGTCTGTAACACTATCTGATGAGAACT 6162
QY 1000 ACCCTACAGTACCGCAGTATTGTTATTAACAGATGATTCAGTGTGATCTTGACATGGTGG 1059
DB 6163 ATCCGACAGTACGATCTTATCAGATTACTACAGATGATGCTTACTTGATATGTTAG 6222
QY 1060 ACCGGCATGGTGTCTGATACAGCCACTTTTTCGGGGCTAAACTGGAAGTACC 1119
DB 6223 ACCGGACAGTGGCTGTGATGATGCAACCTTCTGCCCCGTACCTTAGAAGTTACC 6282
QY 1120 CAAGAAGCATAGCTATTTTCAGCCAGAGATAAGATCAGCCGTCCCATGCTATACAGA 1179
DB 6283 CGAAAAAATGAGTATAGAGCCCGAATATCCGAGTGGGGTTCCATCAGCATGACAGA 6342
QY 1180 ATACATTACAAATGATTTAGCTGACACTTAAGAAATTTGCAAGCTTAACCAATTC 1239
DB 6343 ACACGCTACAAATGTCCTCATTTGCCCACTAAAGAAATTCACACGTACCGCAGATGC 6402
QY 1240 GAGAAATTAACCTGCTTAGATTGCGCGGCAATTTATGTTGTAAGAAATAGCAT 1299
DB 6403 GTGAACTGCCAACACTGAGCTCAGCCGACATTCATGCGAATGCTTTGCAAAATATGCAT 6462
QY 1300 GCATATGAGTACTGCGATACCTTTCCGATTAACCTTATTCGCTTAACATACAGAACG 1359
DB 6463 GTATATGACAGTATTGGGAGAGATGCTGCGAAGCCAAATTAGATTAACACTGAGTTTG 6522
QY 1360 TTACGCAATATGTCACAAAGCTGAAGAGGCCGGAAGACAGACATTTGTCGAATCTC 1419
DB 6523 TCACCGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6582
QY 1420 ATATATCAAAACCGTTCAGAGATACCAATGATCAATTCATGATGATGATGATGATGATGAT 1479
DB 6583 ATATATTTGTCCTCATTTGCAAGAAATGCTTATGATGATGATGATGATGATGATGATGAT 6642
QY 1480 ATGTCAAGTATCTCCGCGACAGAAATACAGAGAGCCGCTTAAGTGCAGGTTATTC 1539
DB 6643 ACGTGAAGTTACACCGAGCAGCAACACAGAGAAAGAACAGAAAGTCAAGTGTATAC 6702
QY 1540 AGGCTGCAATCCCTTTGCTACCGCTTACCTTTGCGGGATTCATCGGGAATTTAGTCCGTA 1599
DB 6703 AAGCCGAGAACCCCTGCGCACATGCTTACTTATGCGGAGTTTACCGGGAATTTAGTGGTA 6762
QY 1600 GACTGAATGCGGCTCTGCGCAATATTCATCTCTCTGATGATGATGATGATGATGATGATGAT 1659
DB 6763 GGGCTTACGCGCTTCTGCTTCAACCTTACACGCTTTTGAACATGTCGGCGAGATTT 6822
QY 1660 TTGATGCGATTTATGCTGAACATTTTCCACACGCGGCAACCGATTTGGAACGACATCG 1719
DB 6823 TTGATGCAATCATAGAGAAACATTTACAGCAAGCCGCGGATCTGGAACCGATATCG 6882
QY 1720 GGTGCTTGAATAAGAGAGACGAGCTATTCGCAATTTGGCGTTGATATCTTGAGG 1779
DB 6883 CATCATTTGCAAAAGCCAGAGAGCGCTATGCGTTAACCGGTCTGATATCTTGAGG 6942
QY 1780 ACTTAGGTGCAACCAACCGCTCTTGAATTTGATGATGAGGCGGGTTCGGAATATCAT 1839
DB 6943 ACCGGGTGATGATCAACCACTGATGATGATGATGATGATGATGATGATGATGATGAT 7002
QY 1840 CTGTGCACTTACAGAGAGAGGTTTAAATTTGTTGTCATGATGATGATGATGATGATGAT 1899
DB 7003 CCAACCATCTACTAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7062
QY 1900 TCTTAACGCTGTTTGTCAACACACTATGATGATGATGATGATGATGATGATGATGATGAT 1959
DB 7063 TCTTCACTATTTTGTCAACACAGTTTGAATGATGATGATGATGATGATGATGATGATGAT 7122
QY 1960 AAGGTTAACAGGTAGAGTGGCGGCTCTATGCGCGAGATGATGATGATGATGATGATGAT 2019
DB 7123 AGCGGCTTAAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7182
QY 2020 TCGTCTCCGACACTTGTATGCGGAGATGCGCACTTGGCTGAACATGGAAGTAAAA 2079

Db	7183	TAGTATCTGACAAAGAAATGCTGAGAGGTGGCCACTGGCTCAATGAGGTTAAGA	7242
Qy	2080	TTATTGATGCAAGTATTTGGTATCAAGAACCCTACTTCTGTGGGGATTTATCTGTGTGG	2139
Db	7243	TCATCGAGCGAGTCAATCGGTGAGAACCACTTACTTCTGGGGCGGATTTTATCTTGAAG	7302
Qy	2140	ACCAAGTAAACGGGCAACGCTGCGAGAGTCCGACACCCTTAATAAGGCTTTTAAAGCTG	2199
Db	7303	ATTGGTACTTCCACAGCTGGCCGCGTGGCGGATCCCTGAAAAGGCTGTTTAAAGTTGG	7362
Qy	2200	GAAAACCATTTGCGAGTGCATGATATACCAGAATGGCAACCGCCGCGGGCACTGCATGATG	2259
Db	7363	GTAATCCGCTCCCAACCCCAACGACGACGAAGAAAGAAAGAGCCGCTCTGTAGATGG	7422
Qy	2260	AAGCAATGCGATGAGAACAGAAATTTGGAATTAACGAGACGATTTAGTGAAGGCGCTGAATATCA	2319
Db	7423	AAACAAAGCGCTGTTTGAAGTAGTATTAACAGGCACTTTAGCACTGGCCGTGACGACCC	7482
Qy	2320	GATACGAGATCATACTGGCAGGCGCTGATCATCAGTCTGTGTCCACGTTAGCCGAAAGCG	2379
Db	7483	GGTATGAGGTAGACAAATATTAACACTGTCTCACTAGGCAATTGGAACTTTCCGACGAGCA	7542
Qy	2380	TTAAGAACTTCAAGAGCAATTAAGGAGGAGCCCAATACCCCTTAGCGGCTGACCTTAATATGG	2439
Db	7543	AAAGGCAATTCACAGCCCACTAGAGGGGAAATTAAGCATCTCAACGCTGTCTTAAATAGT	7602
Qy	2440	TGAGGTAGTACACGCACTTAC-----CCACCGGCGAAGATGTTTCAATACC	2486
Db	7603	CAGCATATGATATTTGATCTGATCTAATATCTAACACACACACACATGATAGAGATTTCT	7662
Qy	2487	CTCAGCTAACTTTTCCACACAGTTTACCTTAACAAATCCGATGGCTTACCGAATCCAAACC	2546
Db	7663	TTTAACTATCTGGCGCGCCGCTTTCCCGGCCCACTGGCATGTGAGGCGCGGAGAA	7722
Qy	2547	CTCTAGGCGCGCGCTGAGGCGGTTTGGGCCCCGCTGGCTGCTCAATTCAGAGATCTTA	2606
Db	7723	GGAGGCAAGCGGGCCCCGATGCTCTGGCCGCAACGGGCTTGGCTTCTCAATCCAGCACTGA	7782
Qy	2607	GGAGGTGATATGTCAACTTTGAAACAACGATCACTTAATCCGCGCCAGGTCCA-	2665
Db	7783	CCACAGCGGTGATGCGCTTACTGATGTGACAGGCACTAAGACCTCAACCCCAAGTCCAC	7842
Qy	2666	-----CGCCCAAGAAAGAAAGAGAGTCTCTTAAGCCAAACCTAATAGCTTTAAAAAG	2720
Db	7843	GCCCCGCCACCCCGCCAGAAAGAGAGCGCCCAAGCAACACCGAAGCCGAAGAAACCA	7902
Qy	2721	AGAGGAGCAAGCCAAAGAGGAGAAAGCC---AAGCTTAATCCAGGAGAAAGCAAGCTA	2777
Db	7903	AAAGCGAGAGAAAGAAAGAAAGACCACTTGCAAAAACCCAAACCGGAAAGAGACAGCGCA	7962
Qy	2778	TGTGTATGAAGTTGGAGTCCGACA	2801
Db	7963	TGGCATTAAGTTGAGGCCCAACA	7986

RESULT 11
 US-09-507-362-101
 : Sequence 101, Application US/09507362
 : Publication No. US20030096397A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Dubensky Jr., Thomas W.
 :
 : Polc, John M.
 :
 : Belli, Barbara A.
 :
 : Schlesinger, Sondra
 :
 : Dryga, Sergey A.
 :
 : Prolov, Ilya
 :
 : TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
 : WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
 : SYNTHESIS
 :
 : NUMBER OF SEQUENCES: 125
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: Seed Intellectual Property Law Group PLLC
 :
 : STREET: 701 Fifth Avenue, Suite 6300
 :
 :
 :

? CITY: Seattle
 ? STATE: Washington
 ? COUNTRY: USA
 ? ZIP: 98104-7092
 ?
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: PatentIn Release #1.0, Version #1.30
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/09/507,362
 ? FILING DATE: 18-Feb-2000
 ? CLASSIFICATION: <Unknown>
 ?
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: McMasters, David D.
 ? REGISTRATION NUMBER: 33,963
 ? REFERENCE/DOCKET NUMBER: 930049.457D6 /1196.011
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (206) 622-4900
 ? TELEFAX: (206) 682-6031
 ?
 ? INFORMATION FOR SEQ ID NO: 101:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 8000 base pairs
 ? TYPE: nucleic acid
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ?
 ? SEQUENCE DESCRIPTION: SEQ ID NO: 101:
 ? US-09-507-362-101

	Query Match	12.7%	Score 824.4; DB 11;	Length 8000;
	Best Local Similarity	61.9%;	Pred. No. 5.4e-235;	
	Matches 1388;	Conservative	0; Mismatches 831;	Indels 25; Gaps 4.
QY	580	ACTGACGGTATGAAACGGGAGCCGTATATTTTCTCATCGGAAACAGCCCAAGTCACTTC	639	
DB	5746	ACTGACTTAACCGGGGGTAGTGGGTACATATTTTTCGACGGACAGCCCTGGCACTTGC	5805	
QY	640	AACGAATTCAGTACGTCAATGTMAAATACTACAGAAACCTATATTGGATCGGGCCGTCCATG	699	
DB	5806	AAAAGAAAGTCGGTCTCTGAGAACCAAGCTTACAGAACCGAAGCTTGAAGACAAATGCTCTGG	5865	
QY	700	AGAAGTATTACGCCCCCGCCTCGATCTCGAAAGAGAAATATTTACAGAAAGAACTGC	759	
DB	5866	AAAGAATTCATGCCCCCGGTGTGTGACACGTGGAAGAGAAACATCTAAATCAGATAC	5922	
QY	760	AATTATGGCCTCTGAAGAAATATGAAGACGTATCAATCAGAAAGTAGAAATATATGA	819	
DB	5926	AGATGATGCCCAACGGAAACCAAAAGTAGGTACCAAGTCTCTGTAAGTAGAATAATATCAGA	5985	
QY	820	AAGCAATTTACAGGGAGAGCACTCATTTCTGGATTGGGCAATATCTATCATCAAGAAATGA	879	
DB	5986	AAGCCATTAACACTGAGGCACTACTGTGAGACTACGACTGTATTAATCTTGCCACAGATC	6045	
QY	880	ATCTCTCGAGTGTTCAGAGTCAATATCTCTGACCAATCTACTCGTCAACGGTAAATTA	939	
DB	6046	AGCCAAAA--GAATGTTATATGATCACTATTCGGAACCAATGTATCTCCAGTAGGTTAACGG	6102	
QY	940	ACAGGTTTACATCTGCAGAGGTGCGGGTTAAAACGTGCACCTTATGTTATCCAAAGAAATT	999	
DB	6103	CGAATCTACTCGAATCCACAGTTCCTCTGAGCTGTCTATACAACTATCTGTGATAGAACT	6162	
QY	1000	ACCTTACAGTAGACAGATTATTTGTATTAACAGATGAATTCAGATGCGTATCTGACATGGTG	1059	
DB	6163	ATTCGACAGTAGAATCTTATACAGTTTACTGACGAGTAGATGCTTACTTGGATATTGGTAG	6222	
QY	1060	ACGGCGCATCGTGTCTGTAGATACAGCCACTTTTGTCTCGGCTAAAATGGAAGCTAAC	1119	
DB	6223	ACGGGACAGTCGCTGCTGGATCTGCAACCTTCTGCCCGCTAAAGCTTGAAGTTAAC	6282	
QY	1120	CAAAAGAGCAATAGTATTTTGACAGCCAAAGATATAATCAGCCCTCCCATCCGCTTATACGA	1179	
DB	6283	CGAAAAAACATGAGATATAGACCCCGGAATATCCGCAATGGGGTTCATATCAGCATGACGA	6342	

QY 700 AGAAGTATTAGCCCGGCTCGATCTCGAAGAGAGAAAATGTTACAGAAAGAACTGC 759
 Db 6758 AAAAGATTTCATGCCCCGGGTGCTCGACAGCTCGAAGAGAGAACTCAAACTCAGATACC 6817
 QY 760 AATTATGGCCCTCTGAAAGAAATAGAGAGATATCAATCAAGAAAGTAAATATAGA 819
 Db 6818 AGATATGCCCAACCAAGCAAAAGAGTACCAAGTCTGTAAGTAAAGTAAATATAGA 6877
 QY 820 AAGCAATTACAGCGAGGAGCTCATTTTCGATTTGGGACATCTATCATCAGAGTGA 879
 Db 6878 AAGCATACCACTAGGAGCTACTGTGAGAGCTAGACTGTATTAATCTTCGACAG--- 6934
 QY 880 ATCTGTGAGTGTATACAGATCAATATCTGTACCAATCTACTCTGTCAACGATTA 939
 Db 6935 ATCAGCCAGAAATGCTATAGATCACTATCCGAAACATTTGATCTCAGATACGACGG 6994
 QY 940 ACAGGTTACATCTGCAAGAGCTGGGTTAAACGTCGAATTAATGTTATCCAGAGATT 999
 Db 6995 CGAATCTCCGATCCAGATTCAGATTCGCTGTAGCTGTATCAACATATCTGCAATGAACT 7054
 QY 1000 ACCCTACAGTAGCCAGTATTATGTATACAGATGAATACGATCGTATCTTGACATGTGG 1059
 Db 7055 ATCCGACAGTAGCATCTTATTCAGATTTACTGACAGATGATGATCTTATCTTGATGTAG 7114
 QY 1060 ACAGGCGATCTGTCTGATAGTACAGCCATTTTGTCCGGCTAACTGAGAGCTACC 1119
 Db 7115 ACAGACAGTGCATGCTGATGATCTGCAACCTTCTGCCCCGCTAAGCTTAGAATTTACC 7174
 QY 1120 CAAAGAACATAGCTATTGTCAGCCAGATTAAGTCAAGCCGTCCCATTCGCTTATACGA 1179
 Db 7175 CGAAAAACATGATAGATAGAGCCCGCAATATCCGAGTGCCTGATCATCAGAGATGACGA 7234
 QY 1180 ATACATCTCAAAATATATTTGGCTGAGCTATTAAGAAATGCAACGTTACCCAAATGC 1239
 Db 7235 ACAGCTCAAAATATGCTCATTTGCCGCACTTAAGAAATTTGCAACGTCAGCGAGATGC 7294
 QY 1240 GAGAAATTACTGCTTATGATTCGGGCGCATTTAATGTTGATTTTCAAGAAATAGCAT 1299
 Db 7295 GTGAATCTCCCAACATGAGACTCAGGACATTCATATGTGAATGCTTTGAAAAATATGAT 7354
 QY 1300 GCAATGATGATGATCTGAGTACCTTTCCGATTAACCTTATTTGGCTTAACTACAGAAAG 1359
 Db 7355 GTAAAGACAGATGATTTGGAGAGGATTCGCTCGGAAACCAATTAAGATTACACTGAGTTTG 7414
 QY 1360 TTAAGCAATATGTGCAAAAGCTGAAAGGCGGAAAGCAGCAGCTTTGTTGGAAATACTC 1419
 Db 7415 TCACCGCATATGTAGCTAGCTAGTAAAGGCCCTTAAGGCGCGCACATTTTGGAAAGAGCT 7474
 QY 1420 ATAAATCTAAAAACGTTGCAAGAGATACCAATGATCAATTCGTCAATGATCTAAAGAGAG 1479
 Db 7475 ATAAATTTGGTCCCATTTGCAAGAGATGCTATGATGATTTGCTCATGACATGAAAGAGAG 7534
 QY 1480 ATGTCAAAATTAATCTCCGCGACGAAACATACAGAGAGCGGCTTAAGGTGACGTTATTC 1539
 Db 7535 ACGTAAAGATTACACAGGCAAGAAACACAGAAAGAAAGCCGAAATACAAAGTATAC 7594
 QY 1540 AGGCTGAGATCCCTTGTCTACCGCTTACCTTTGGGGATTCATGCGGAATTAAGCCGTA 1599
 Db 7595 AAGCCGCAAAACCTTGGGAGCTGTACTTATATGCGGATTTCAACGGGAATTAAGCGCTA 7654
 QY 1600 GACTGAATGCGGTCTTGTGCAAAATATCAATCTCTTGCACATGTCAAGGGAAGATT 1659
 Db 7655 GGCCTTAAGCGCGTCTTGTCTTCAAAATCAACACCTTTTGAACATGTCCGCGGAGATT 7714
 QY 1660 TTGATGCAATTAATGCTGAACATTTCCACACGCGGACCCAGATTAATGGAAGGAGATCG 1719
 Db 7715 TTGATGCAATTAATGCAAAACATTTCAAGCAAGGCGGACCTGGAATCTGAGAGCGATATCG 7774
 QY 1720 GGTGTTGATTAAGGAAGGAGCGGATTCGCAATTTGGCGGTTGATGATCTTGAGG 1779
 Db 7775 CATCTTTGACAAAGGCAAGGCGGATTCGCTTAACCGGCTGATGATCTTGAGG 7834
 QY 1780 ACTTAGTGTGACCAACCGCTCTTGAATTTGATGAGGCGGCTTTCGCAATATCACAT 1839

Db 7835 ACTGCGTGTGATCAACACACTACGACTTGATGATGAGCGCTTTGAGAAATATCAT 7894
 QY 1840 CTGTGACACTTACCTACAGAAACAGAGTTAAATTTGGTCCATGATGAAATCCGATCT 1899
 Db 7895 CCACCATCTACTACGGGTACTCGTTTAAATTCGGGCGATGATGAAATCCGAAATCT 7954
 QY 1900 TCTTAACGCTGTTTGTCAACACATGATCAATATCATATGCTGTGACAGATACCTG 1959
 Db 7955 TCTTCAACTTTTGTCAACACATGTTGATGATGATGATGATGATGATGATGATGATG 8014
 QY 1960 AACGTTAACACGTCAGCGTCCGCGCTCTATCGCGAGATTAATGATGATGATGATG 2019
 Db 8015 AGCGCTTAAACGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 8074
 QY 2020 TCGTCTCGACACTTGAAGCGGAGATGATGATGATGATGATGATGATGATGATGATG 2079
 Db 8075 TAGATCTGACAAAGAAATGCTGAGAGTGGCCCACTGCTCAACATGAGATTAAGA 8134
 QY 2080 TTATGATGACATTAATGATCAAAAGCACTTCTGTGAGGAGTTTATCTGATG 2139
 Db 8135 TCATGAGCAGATCATCGGTAGAGACCACTTACTTCTGCGCGGATTTATCTTGCAAG 8194
 QY 2140 ACCAGATTAACAGGACAGCTTCAGAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
 Db 8195 ATTCGTTACTTCCACAGACGTCGCGGAGATCCCTGAAAGGCTGTTAAGTTGG 8254
 QY 2200 GAAAACTTTCGCTGATGATTAACCAAGTGCAGACCGCGCGGCACTGCAATGATG 2259
 Db 8255 GTAAACCGCTCCAGCCGACAGAGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 8314
 QY 2260 AAGCAATGATGAGAAAGAAATGGAATTTAGAGAGAGTGAAGGCGGAGAAATCA 2319
 Db 8315 AAACAAAGCGGTGTTTGAAGTATGATTAACAGGACCTTATGACAGTGCCTGAGAGACC 8374
 QY 2320 GATACGATCATCTGCGAGGCTGATCATCACTCTCTGTCCAGCTTAAGCCGAAAGCG 2379
 Db 8375 GGTATGAGTGAACAATTTACACCTGCTTACTGAGCATTTGAGAACTTTGCGCAGAGCA 8434
 QY 2380 TTAAGAACTTCAAGAGCTTAAGAGAGAGCCCAATCACTCTTAAGGCTGATTAATG 2439
 Db 8435 AAAGAGATTCGAAAGCCATCAGAGGGAATTAAGCATCTTACGCTGCTTAATATGT 8494
 QY 2440 TGACGTAGTAGA 2451
 Db 8495 CAGCATAGTAGA 8506

RESULT 13
 US-09-275-883-1
 ; Sequence 1, Application US/09275883
 ; Publication No. US20030053988A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Renner, Wolfgang A.
 ; APPLICANT: Nieba, Lars
 ; APPLICANT: Boersma, Marco
 ; TITLE OF INVENTION: Inducible Alphaviral Gene Expression System
 ; FILE REFERENCE: 1700.0020001
 ; CURRENT APPLICATION NUMBER: US/09/275.883
 ; EARLIER FILING DATE: 1999-03-25
 ; EARLIER FILING DATE: 1998-03-27
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 11282
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:cDNA
 US-09-275-883-1
 Query Match 12.6%; Score 816; DB 11; Length 11282;

SEQ ID NO 4
LENGTH: 8100
TYPE: DNA
ORGANISM: Semliki Forest virus
US-09-190-246-4

Query Match 11.8%; Score 764; DB 12; Length 8100;
Best Local Similarity 63.2%; Pred. No. 7.2e-217;
Matches 1175; Conservative 0; Mismatches 685; Indels 0; Gaps 0;

QY 594 GCGGAGCGTATATTTCTCATCGGAAACAGGCCAAGTCACTTCAACAGAAATCATGTA 653
Db 5251 GCGGGTGATATATTTCTCTCCGCGACACTGGCAGCGGACATTTTCAACAAATATCCGTT 5310
QY 654 GGTCAATGTAACATCAAGAAACCTATATTGATCGGGCGCTCCAGAGAAATATACGCC 713
Db 5311 AGGCGCAAAATCTTCAGTGGCCACAACTGATGGGTTCAGAGAGAGAAATGATACCG 5370
QY 714 CCGGCGCTCATCTTCGAAAGAGAGAAAATGTTACAGAAAGAAAGTCAATTAATGCGCTCT 773
Db 5371 CCAAAATTTGATATCTGAGAGGAGAGAGAGTTCGCTGAGAAATGCAATGCAATGCAATCG 5430
QY 774 GAAGAAATAGAGAGATATCATACGAAAGTGAATAATGAAAGCAATTAACAGCG 833
Db 5431 GAGGCTAATTAAGATGATACCAAGTCTGCAAAAGTGAGAAATGAAAGCCACGCGTGGTG 5490
QY 834 GAGGAGCTCATTTCTGAGTTGGGCAATATCATGAGAAATGTAATCTGTGAGTGT 893
Db 5491 GACAGGCTCAATCGGGGGCCAGATTGTACAGGGAGGAGGATGAGCGCGCATACCAACA 5550
QY 894 TACAGAGTCAATTAATCTGTATACCAATCTACTGTCAGAGTAAATTAAGGTTTACATCT 953
Db 5551 TACGGGTTGCGGATCCCCCGCCGCTACTCTCCCTACCTGATGAGAAAGATTTCCAAGC 5610
QY 954 GCAGAGTGGCGGTTAAACCTGCACTTAATTCAGAGAAATTAATCTTACAGTACC 1013
Db 5611 CCCGATGAGCAATCGCAGCGGCAACCAATCACTATCCAGAAATTAATCCCAAGTGCGG 5670
QY 1014 AGTTATGTATTAACGATGAATAGATGCGTATCTTGAATGCTGAGAGCGGCGCATCGTC 1073
Db 5671 TCGTACAGATTAACGATGAATAGATGCGTATCTTGAATGCTGAGAGCGGCGCATCGTC 5730
QY 1074 TGTCTAGATACAGCACTTTTGTCCGCTAACTGAGAGAGTCCCAAGAGATAGC 1133
Db 5731 TGTCTGAGACAGAGCACTTCTGCCCCGAGAGCTCCGCTGATCCCAAGATCATATCG 5790
QY 1134 TATTTGACGCAAGATTAAGATCAGCGCTCCATGCTTATACAGATTAATCAATCA 1193
Db 5791 TACCAACGAGCGGACTGTAGCGAGTCCGCTCCGCTACCTTTCAAGAACACTACAGAAC 5850
QY 1194 GTATTGGCTGACGATTAAGAAAGAAATGCAACGTTACCCAAATGCGAGATTAATCTGTC 1253
Db 5851 GTGCTAGGCGCGCCACCAAGAGAAATGCAACGTTACCCAAATGCGAGATTAATCTGTC 5910
QY 1254 TTAGATTGGGCGCATTTAATGTTGATTTTCAAGAAATAGCATGCAATGATAGTAC 1313
Db 5911 ATGACTGCGGAGTTCATAGTGAAGTCTTCAAGCCCTATGCTGCTCCGAGAAATAT 5970
QY 1314 TGGGATACCTTTCCGATTAACCTTATTTGGCTTAATACAGAGAAAGTTCAGCAATATGTG 1373
Db 5971 TGGGAAAGAAATGCTTAACCACTTATCCGGAATTAACCACTGAGAAATCACTATATGTG 6030
QY 1374 ACAAGCTGAAGAGGCGCAAGAGCAAGATTTGTTGGGAATATCTAATTAACCG 1433
Db 6031 ACCAAATTAAGAGGCGCCAAAGCTGCTCTTGTTCGTTAAGAACCAACCTTGTGTTCCG 6090
QY 1434 TTGCGAGATACCAATGATCAATTTCTCATGATCTTAAGAGAGATGTCAAAGTTACT 1493
Db 6091 CTGCGAGAGGTTCCCATGAGATTCAGGCTGCAATGAAAGAGATGTCAAAGTCACT 6150
QY 1494 CCCGCGACGAAACATACAGAGAGCGGCTTAAGGTGAGGTTATTCAGGCTGCAAGATCC 1553
Db 6151 CCAAGGAGGAAACACACAGAGAGAAAGACCAAAAGTCCAGTAAATTCAGAGCGGAGCCA 6210

QY 1554 CTGCTACCGCTTACCTTTGCGGAGATCCAGGAAATTAATCCGTAAGCTGAATCGGNG 1613
Db 6211 TTGGCGACCGCTTACCTGTGCGGATCCACAGGAAATTAAGAGAGACTAAATCTGTG 6270
QY 1614 CTCTGCCAAATATCATCTCTCTGCAATGTCAGCGGAGATTTTGAATGATGCAATTT 1673
Db 6271 TTAGGCCCTAAGTGAACATATGTTGATATGTCGCGGAAAGCTTTGAGGCAATCATC 6330
QY 1674 GCTGAACATTTCCACACGCGGACCCAGTATGGAAGAGCAATCGCTGTTGATATA 1733
Db 6331 GCTCTCACTTCCACACGAGAGACCGGTTCTAAGAGAGGCAATGCAATTCAGACAAA 6390
QY 1734 AGCGAAGCGAGCTATCGCAATTTGCGCGTGAATGATCCCTGAGAACTTAAGGTGAC 1793
Db 6391 AGCCAGAGAGCTCTTGTGCTTACAGTTTAATGATCTCGAAGATCTAGGGGTGAT 6450
QY 1794 CAACCGCTCTTGAATTTGATAGAGCGCGCTTCGCAATATCAATCTGTCACCTTACT 1853
Db 6451 CAGTACCTGCTGGACTTGATGAGGAGAGCTTTGGGAAATATCAAGCTGTCACTTACA 6510
QY 1854 ACAGAAACGAGTTTAAATTTGTCATGATGAATCCGATATGTTCTTAACCTGTTT 1913
Db 6511 ACTGCAAGCGCTTCAAGTTGAGCTATGATGAATCGGCAATGTTCTGACTTGTGTT 6570
QY 1914 GTCAACACACTAGTCATATGATGATGCTAGAGAGTACGATCAAGTGAACGGTTAACCA 1973
Db 6571 ATTAACTGTTTGAACATTAATCAATGACAGAGGATCTGAGACAGACTCATGAC 6630
QY 1974 TCAGCTGCGCGGCTTATGCGGAGATTAACATGATGATGATGTCGTCGACACC 2033
Db 6631 TCGGCTGTGCGGCTTATGCGGAGAGCAACATCGTTCAGGAGTATCTCCGACAG 6690
QY 2034 TTGATGGGAGAGATGCGCACTTGCTGAACATGGAAGTAAATTAATGATGCACTT 2093
Db 6691 CTGATGGGAGAGATGCGCGCTCGTGGGTCAACATGAGGTGAAGATCATTTGACGCTGTC 6750
QY 2094 ATTGATTAACAGAACCTTATCTGAGGGGATTTATCCGTGAGCAATTAACAGGC 2153
Db 6751 ATGGCGAAAAACCCCAATATTTTGTGGGGATTCATAGTTTGAAGCGTCAACAG 6810
QY 2154 ACAGCTGCAAGATGCGAGACCTCTTAAAGGCTTTTAAGTTGAGTGAACCAATGCGCA 2213
Db 6811 ACCGCTGCGCTGTTTGAAGACCACTTAAGGCGCTTCAAGTTGGGTAACCGCTAACA 6870
QY 2214 GTCATGATACCAAGACTGCGACCGCGCGGCACTGATGATGAAGCAATGCGATGG 2273
Db 6871 GCTGAAGACAAAGCAGAGCAAGAGACAGGAGACACTGAGTGAAGGATTAAGCAAGTGG 6930
QY 2274 AACGAAATGGAATTAAGGAGAGATTAAGAGGCGGTGAATCCAGATACGATATCTA 2333
Db 6931 TTCCGACAGGCTTGGGCGCGGAATGAGGTGCACTTAATCATGATTAAGAGTGAAG 6990
QY 2334 CTGCGAGGCTGATCATCAGCTCTGTCCAGTTTACCGGAAACGTTAAGAACTTCAAG 2393
Db 6991 GCGTCAAAAGTATCTCATAGCCATGAGCACTTTGGCGAGGAGCATTTAAGCGTTTAA 7050
QY 2394 AGCATTAAGAGGAGCCCAATCACCTCTACGCGTGAACCTTAATAGTGAAGTGAAGCA 2453
Db 7051 AAATTAAGAGGAGCTGTTATTAACCTCTACGCGGCTTGAATGTTGCGTTAATACCA 7110

RESULT 15
US-09-994-412-3
; Sequence 3, Application US/09994412
; Publication No. US20030059943A1
; GENERAL INFORMATION:
; APPLICANT: Cerna, Ulrich
; APPLICANT: Lundstrom, Kenneth
; TITLE OF INVENTION: Inhibition of Expression of a Target Gene
; FILE REFERENCE: 20787
; CURRENT APPLICATION NUMBER: US/09/994,412
; CURRENT FILING DATE: 2002-03-12

NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 10610
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pSFV2gen(pD) vector
US-09-994-412-3

Query Match 11.8%; Score 764; DB 11; Length 10610;
Best Local Similarity 63.2%; Pred. No. 8.8e-217;
Matches 1175; Conservative 0; Mismatches 685; Indels 0; Gaps 0;

QY 594 GCGGACGCTATATTTCTCATCGGAAACGGCCAGGTCCCTTCAACGAAATTCGTA 653
DB GCGGCTCATATATTTCTCTCGACACTGCGACGGACATTTACAAACAAAATTCGGT 5588
QY 654 CGTCAATGTAACTACAAAGAACTATATTGGATCGGCGCTCCATGAGAAATTTACGCC 713
DB AGGACAGCAATCTCCAGTGGCCACAACTGATCGCGTCCAGAGAGAAATTTATCCCG 5648
QY 714 CCGGCTCGATCTCGAAGAGAAATGTTACAGAAATTCGCAATTATGCGCTCT 773
DB CCAAAATTTGATCTGAGAGGAGAAAGCTGTGCTGAAATGACAGATCACCCATCG 5708
QY 774 GAAGAAATAGAACAGGTATCAATCAAGAAATAGAAATATGAAAGCAATTCACGC 833
DB GAGGCTATATAGATTCGATACCACTCTCGAAAGTGAAGAAATGAAAGCCAGGTGGTG 5768
QY 834 GAGGACCTATTTCTGATTTGGGACATATCTATCATCAGAAATTCCTGCGAGTGT 893
DB GACAGCTCATATGCGGGGCCAGATTGACAGGAGGACGTAAGCCGATACCAACA 5828
QY 894 TACAGAGTCAATTTCTGTACCAATCTACTGTCAAGGTAATTAAGGTTTACATCT 953
DB TACCGGTTCCGTTACCCCGCCCGTGTACTCCCTACCGTATCGAAAGATTCGCAAGC 5888
QY 954 GCAAGGTCCGGCTTAAACGTCGAACTTACTTACCAAGAAATTCCTACGATGCC 1013
DB CCCGATGAGCAATTCGACGCTGCAAGCAATTCCTATCCAGAAATTTACCCAAAGTGGCG 5948
QY 1014 AGTATTGTATTAAGATGAATGAATGATGATCTTTCGATGTCGAGCGCATTCGTC 1073
DB TCGTACCAAGTATTAAGATGAATGATGATGATGATGATGATGATGATGATGATGAT 6008
QY 1074 TGTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133
DB TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6068
QY 1134 TATTTGACGCGAGATTAAGATCAAGCCCTCCATCGCTTATACGAATTCATCAAAAT 1193
DB TACCAACGACGCTGATGAGATGAGTCCGCTCCGCTACCTTTCCAGAACACTACAGAAC 6128
QY 1194 GTATTGCTGACGCTATCAAAAGAAATGCAAGCTTACCCAAATTCGCAATTCCTGTC 1253
DB GTGCTAGGCGGTGACCAAGAAATGCAAGCTTACGCAAGTTCGCAAGTTCGCAAGTTC 6188
QY 1254 TTAATTCGCGGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
DB ATGATCTGCGAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6248
QY 1314 TGGGATACCTTTCCGATTAACCTTATTTGGCTTAACTACAGAAAGTTCAGCAATATG 1373
DB TGGGAAATATGCTAAACCACTATCGGATTAACCACTGAGAAATATCACTATATG 6308
QY 1374 ACAAGCTGAAGGCGCAAGACAGACATTTGTTGGAATTAATCATATCTAAACCG 1433
DB ACCAAATTTGAAGGCGCAAGACAGACATTTGTTGGAATTAATCATATCTAAACCG 6368
QY 1434 TTGAGAGATACCAATGATCAATTCGTATGATCTTAAAGAGATGTCAAAGTTACT 1493
DB CTGACAGAGGTTCCATGACAGATTCAGGTCGATGAACGAGATGTCAAAGTCACT 6428

QY 1494 CCGGACGAAACATACAGAGAGCCGCTAAGGTGAGTTATTCAGGCTGCGATCCC 1553
DB CAGAGAGCAAAACACACAGAGAAAGACCCAAAGTCCAGGTAATTCAGAGCGGAGCCA 6488
QY 1554 CTGCTACCGCTTACCTTTGCGGAGTCCATCGGAAATTAATTCGATCAATTCGCTG 1613
DB TTGCGCAGCCGCTTACCTGTCGCGATCCACAGGAAATTAATGAAGACATAATGCTGTG 6548
QY 1614 CTTCGCAATATTCATCTCTTCGATATGTCAGCGGAAATTTGATGCGATTAAT 1673
DB TTACGCTTACGTCAGACATGTTGATGATGATGATGATGATGATGATGATGATGATG 6608
QY 1674 GCTGAACATTTCCACGACGCGACCCAGTATTTGAAACGACATCGCGTGTGATATAA 1733
DB GCTCTACCTTCCACCGAGAGACCGGTTCTAGAGCGACATGATCATTCGACAAA 6668
QY 1734 AGCAGAGACGATCGCTATCGCATTTTCGCGTGTATGATTCCTTGAAGACTTAACTG 1793
DB AGCAGAGACGATCGCTATCGCATTTTCGCGTGTATGATTCCTTGAAGACTTAACTG 6728
QY 1794 CAACCGCTCTTAATTTGATGAGAGCGCGGCTTGGCAATATCATCTGTCACCTACT 1853
DB CAGTACCTGCTGACCTTGAATGAGGACGCTTTGGGAAATATTCAGCTGTACCTACCA 6788
QY 1854 ACAGGAACGAGTTTAAATTTGGTGCATGATGAAATCCGCTATGTTCTTAAAGCTGTT 1913
DB ACTGACGCGGCTTCAGATTCGAGATGATGATGATGATGATGATGATGATGATGATG 6788
QY 1914 GTCAACCACTAGTCAATATATGATTTGCTAGCAGAGTACTAGTGAGACGTTAACAC 1973
DB ATTAACTGCTTTTGAACATACATATGACAGGAGGTTACTGAGCAGAGATCTACATG 6908
QY 1974 TCAAGTCCGCGCTTATTCGCGAGATTAATGATGATGATGATGATGATGATGATGATG 2033
DB TCCGCTGTCGCGCTTATTCGCGAGATTAATGATGATGATGATGATGATGATGATGATG 6968
QY 2034 TTGATGCGGAGAGATGCGGCACTTGGCTGAACATGGAATTAATTTATGTCAGTT 2093
DB CTGATGCGGAGAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7028
QY 2094 ATTGATTAACAGCCTTACTTCTGTGGGGAATTAATCCTGTGACCAAGATTAACAGGC 2153
DB ATGGCGAAACCCCATATTTTGTGGGGAATTAATGATTTTGAACAGCTACACAG 7088
QY 2154 ACAGCTTCGAGATTCGAGACCTCTTAAAGCTTTTAAAGCTTGAAGAAACCTTACCA 2213
DB ACCGCTGCGGCTTGTGAGACCACTTAAAGCTTGTCAAGTGGGTAAGCCGCTTAA 7148
QY 2214 GTGATGATTCGCAAGCTGCGACCGCGCGCGGCACTGATGATGAAGCAATGCGATGG 2273
DB GCTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7208
QY 2274 AACGAATTTGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2333
DB TTCCGAGACAGGCTTGGGAGCGAACTGAGAGTGGCACTTAACTATGATGATGATGAG 7268
QY 2334 CTGCGAGGCTGATCATCATCTCTCTGACAGTTCAGCGAAGAGCTTAAAGAACTTCA 2393
DB GCTGCAAAAGTATCTCATAGCATGCGACCTTGGCGAGAGACATTAAGGCGCTTAA 7328
QY 2394 AGCATAGAGAGAGCCCAATCACCGCTTACGCGTGAACCTTAATGATGATGATGATG 2453
DB AAATTTGAGAGAGCTGTTATTAACCTTACGCGGCTCTGATGATGATGATGATGATG 7388

Search completed: November 16, 2003, 14:53:10
Job time : 1807 secs

Query Match 90.7%; Score 5879.2; DB 3; Length 11492;
Best Local Similarity 96.2%; Pred. No. 0; Mismatches 220; Indels 29; Gaps 20;
Matches 6239; Conservative 0;

QY 1 CACCGCAGCAACGCAACGTAATATGTGAACCAACTGTGCGCAAGATATACCAAGTC 60
DB 5012 CATGCCAACAACGCAACGTAATATGTGAACCAACTGTGCGCAAGATATACCAAGTC 5071
QY 61 CCGGTATTTCTAGACGACCAAGTGGGATCACTGCTATCGCTTGGGGGCTCTCGTTCT 120
DB 5072 CCGGTATTTCTAGACGACCAAGTGGGATCACTGCTATCGCTTGGGGGCTCTCGTTCT 5131
QY 121 CTAGCTGCGACACGGCTCGTCAACGGCTGGAGCGACTATAGACGGCGGTTTGGTTA 180
DB 5132 CTAGCTGCGACACGGCTCGTCAACGGCTGGAGCGACTATAGACGGCGGTTTGGTTA 5191
QY 181 CAGCTGATGTCATCAAGCGAACAGTCTACGTGGAGCATCCCTAGTCTCTGGCTTGG 240
DB 5192 CAGCTGATGTCATCAAGCGAACAGTCTACGTGGAGCATCCCTAGTCTCTGGCTTGG 5251
QY 241 ACGTCCAGCTGCTCTTGAAGCTCACTGATTCCTCACTGGAGTATTCGAATGTCATGAGCT 300
DB 5252 ACGTCCAGCTGCTCTTGAAGCTCACTGATTCCTCACTGGAGTATTCGAATGTCATGAGCT 5311
QY 301 TTGAAGTGAAGAACCACTGTGTACAGGACCTAACCTGGGAGTGGCGAAGCTGTGGAC 360
DB 5312 TTGAAGTGAAGAACCACTGTGTACAGGACCTAACCTGGGAGTGGCGAAGCTGTGGAC 5371
QY 361 TGCGCGAATATATGCAAGACTTCAATATGCTCTCTTTCAGATTCTTTCGAGCTACAGAC 420
DB 5372 TGCGCGAATATATGCAAGACTTCAATATGCTCTCTTTCAGATTCTTTCGAGCTACAGAC 5431
QY 421 CAGTACCGGCAACCGGAGAGCGCCCATCCCTATCACTAGTGAAGCGCTTCCGACCTC 480
DB 5432 CAGTACCGGCAACCGGAGAGCGCCCATCCCTATCACTAGTGAAGCGCTTCCGACCTC 5491
QY 481 CAGTTCGAAAGCCAGGAGACTAAGTACCAACCAACGAGGAGTGGCTGAGAGCATCT 540
DB 5492 CAGTTCGAAAGCCAGGAGACTAAGTACCAACCAACGAGGAGTGGCTGAGAGCATCT 5551
QY 541 CAGAGCGAGTGGAGCGAGTACATCCGTCAACCTCACTGACGATAGAGCGGAG 600
DB 5552 CAGAGCGAGTGGAGCGAGTACATCCGTCAACCTCACTGACGATAGAGCGGAG 5610
QY 601 CGTATATTTTCTCATCGAAAACAGGCCAAGCTCACTTCAACAGAAATCACTAGTCAAT 660
DB 5611 CGTATATTTTCTCATCGAAAACAGGCCAAGCTCACTTCAACAGAAATCACTAGTCAAT 5667
QY 661 GTAAACTACAGAACTATATTTGATCGGGCGCTCAATGAGAAATTAAGCCCGCGGCC 720
DB 5668 GTAAACTACAGAACTATATTTGATCGGGCGCTCAATGAGAAATTAAGCCCGCGGCC 5726
QY 721 TCGATCTCGAAAGAGAGAAATGTTACAGAAAGAACTGCAATTAATGCGCTC-TGAAGGA 779
DB 5727 TCGATCTCGAAAGAGAGAGATGTTACAGAAAGAACTGCAATTAATGCGCTC-TGAAGGA 5786
QY 780 AATGAAAGCAGGTATCAATCAAGAAAGTA-GAAATATGAAAGCAATTACAGCGGAGCG 838
DB 5787 AATGAAAGCAGGTATCAATCAAGAAAGTAAGAAATATGAAAGCAATTACAGCGGAGCG 5846
QY 839 ACTGATTTTGGGATTTGGGCAATATCTATCATGAGAGTGAATCTGTGCGAGTTTCAAG 898
DB 5847 ACTGATTTTGGGATTTGGGCAATATCTATCATGAGAGTGAATCTGTGCGAGTTTCAAG 5906
QY 899 AGTCAATATCTGTACCAATCTACTGTCACACGTAATTAACAGTTTACATCTGAGA 958
DB 5907 AGTCAATATCTGTACCAATCTACTGTCACACGTAATTAAGTTTACATCTGAGA 5966
QY 959 GGTGCGGTTAAAGCGTGAACCTTAGTTATCAAGAGATTAACCTTACAGTACCACTTA 1018
DB 5967 AGTGGCGGTAAAGTTGGAACCTTAGTTATCACTACAGAGATTAACCTTACAGTACCACTTA 6026

QY 1019 TTGTATATACAGATGATATACAGTGGTATCTTGAATGTTGGAACGCGCATCTGCTGCTCT 1078
DB 6027 TTGTATATACAGATGATATACAGTGGTATCTTGAATGTTGGAACGCGCATCTGCTGCTCT 6084
QY 1079 AGATACAGCCACTTTTGTGTCGAGCTAAACGTGAAGTACCCAAAGAGCATATATTT 1138
DB 6085 AGATACAGCCACTTTTGTGTCGAGCTAAACGTGAAGTACCCAAAGAGCATATATTT 6140
QY 1139 GCAGCCAGATATGAATACACCCCTCCCATGCGCTATACAAATATCAATTAAGAAATGTAAT 1198
DB 6141 GCAGCCAGATATGAATACACCCCTCCCATGCGCTATACAAATATCAATTAAGAAATGTAAT 6200
QY 1199 GCGTGCAGTACTATAAAGGAATGCAAGTTCACCAATGCGAATATACCTGTCTAGA 1258
DB 6201 -GCTGCAGCCACTATAAAGGAATGCAAGTTCACCAATGCGAATATACCTGTCTAGT -A 6257
QY 1259 TTGCGCGGCAATTAATGTTGTTCAAGAAATACGATGCAATGATAGTACTGGGA 1318
DB 6258 TTGCGCGGCAATTAATGTTGTTCAAGAAATACGATGCAATGATAGTACTGGGA 6317
QY 1319 TACCTTTGCGGATTAACCTTATTTGGGCTAACTACAGAGAAAGTTACCGAATATGACAAA 1378
DB 6318 TACCTTTGCGGATTAACCTTATTTGGGCTAACTACAGAGAAAGTTACCGAATATGACAAA 6377
QY 1379 GCTGAAAGGCGCGAAAGCAGACAGATTTGTTGCGAATACTCATATATCTAAACCGTTGCA 1438
DB 6378 GCTGAAAGGCGCGAAAGCAGACAGATTTGTTGCGAATACTCATATATCTAAACCGTTGCA 6437
QY 1439 GAGATATCAATGATCAATTCGTATGATCTTAAGAGAGATGTCAAAGTTACTCCCG 1498
DB 6438 GAGATATCAATGATCAATTCGTATGATCTTAAGAGAGATGTCAAAGTTACTCCCG 6496
QY 1499 CAGGAAACATACAGAGAGGCGCTAAGTGGCAGTATTCAGGCTGCAATCCCTTGC 1558
DB 6497 CAGGAAACATACAGAGAGGCGCTAAGTGGCAGTATTCAGGCTGCAATCCCTTGC 6556
QY 1559 TACCGCTTACCTTTGCGGAGTCCATCGGGAATTAAGTCTGAGACTGAGTCTTCT 1618
DB 6557 TACCGCTTACCTTTGCGGAGTCCATCGGGAATTAAGTCTGAGACTGAGTCTTCT 6614
QY 1619 GCCAATATCATACTCTCTTTCAGATGTCAGCGGAAATTTTGTGATGCAATATGCTGA 1678
DB 6615 GCCAATATCATACTCTCTTTCAGATGTCAGCGGAAATTTTGTGATGCAATATGCTGA 6674
QY 1679 ACATTTTCACACAGGCGACCCAGTATTTGGAAGAGACATGCGCTGTTGATTAAGCGA 1738
DB 6675 ACATTTTCACACAGGCGACCCAGTATTTGGAAGAGACATGCGCTGTTGATTAAGCGA 6734
QY 1739 AGAGAGCGTATCGCAATTTGCGGCTGATGATCTTGAAGAGACTTAAGTGTGCAACCAAC 1798
DB 6735 AGAGAGCGTATCGCAATTTGCGGCTGATGATCTTGAAGAGACTTAAGTGTGCAACCAAC 6794
QY 1799 GCTCTTGAATTTGATAGAGCGCGCTTGGCAATATCAATCTGTGCACTTACTACAG 1858
DB 6795 GCTCTTGAATTTGATAGAGCGCGCTTGGCAATATCAATCTGTGCACTTACTACAG 6854
QY 1859 AAGGAGTTTAAATTTGCTGTCATGATGAATCCGGAATTTCTTAAACGCTGTTGCAA 1918
DB 6855 AAGGAGTTTAAATTTGCTGTCATGATGAATCCGGAATTTCTTAAACGCTGTTGCAA 6914
QY 1919 CACACTAGTCAATATCATGATTTGCTAGCAGAGTACTACGTAACGTTAACAGCTGAGC 1978
DB 6915 CACACTAGTCAATATCATGATTTGCTAGCAGAGTACTACGTAACGTTAACAGCTGAGC 6974
QY 1979 GTGCGCGGCTCTATCGCGCAGATTAACATAGTCAATGTGTGCTTCCGACACTTTGAT 2038
DB 6975 GTGCGCGGCTCTATCGCGCAGATTAACATAGTCAATGTGTGCTTCCGACACTTTGAT 7032
QY 2039 GCGGAGAGATGCGGCACTTGGCTGAACATGGAAGTAAATTAATGATGATCACTATGG 2098
DB 7033 GCGGAGAGATGCGGCACTTGGCTGAACATGGAAGTAAATTAATGATGATCACTATGG 7092
QY 2099 TATCAAGACACCTACTTCTGTGGGGGATTTATCTGTGAGCAAGTAAACAGGACAGC 2158

Db	7093	TATCAAGACACCTTACTTCTG-CGGGAGTTCACTCTGGATCGATCAAC-GACACAGC	7156
Oy	2159	CTGCAGATCGCAGACCCCTTAAAAAGCTTTTAAAGTTGGAAAAACATTGCCAGTCGA	2218
Db	7251	CTGTAG-GTCGAGACCCCTCTAAAAAGGCTTTTAAAGTTGGAAAAACATTGCCAGTCGA	7205
Oy	2219	TGATATCCCAAGATCTCGAACCCGCCCGG3CACTGCATGATGAAAGCAATGCGATGGAAACG	2278
Db	7210	TGACACCCCAAGCTGTGACCCGCGCGG3CACTGCATGATGAAAGCAATGCGATGGAAACG	7266
Oy	2279	AATTGSAATTACGGACGATAGTAAAGCCGTAAATTCAGATTCAGATCATCTGGC	2333
Db	7270	AATTGSAATTACGGACGATAGTAAAGCCGTAAATTCAGATTCAGATCATCTGGC	7322
Oy	2339	AGGCGTGTATCAACGTCCTGTGTCCAGCTTAGCCGAAAGCGTTAAGAACTTCAAGAGAT	2396
Db	7330	AGGCGTGTATCAACGTCCTGTGTCCAGCTTAGCCGAAAGCGTTAAGAACTTCAAGAGAT	7385
Oy	2399	AAGAGGAGCCCAATCACCTCTTACGGCTGACTTAAATAGTGACGTAGTAAACAGCAC	2456
Db	7390	AAGAGGAGCCCAATCACCTCTTACGGCTGACTTAAATAGTGACGTAGTAAACAGCAC	7449
Oy	2459	CTACCCACCCGC-AGAAATGTTCCATACCCCTAGCTGAACCTTCCACAGATTACCTTAC	2517
Db	7450	CTACCCACCCGCAGAAAAATGTTTCCATACCCCTAGCTGAACCTTCCACAGATTACCTTAC	7509
Oy	2518	AAATCCGATGGCTTACCGAGATCCAAACCCCTCTAGGGGCGCTGGAAGCCGTTTGCGGC	2577
Db	7510	AAATCCGATGGCTTACCGAGATCCAAACCCCTCTAGGGGCGCTGGAAGCCGTTTGCGGC	7566
Oy	2578	CCCGCTGGCTGCTCAATCGAAGATCTTAGAGAGTGAATAGTCACTTGACTTTCAACA	2637
Db	7570	CCCGCTGGCTGCTCAATCGAAGATCTTAGAGAGTGAATAGTCACTTGACTTTCAACA	7629
Oy	2638	ACGATCACCTAATCCGCGCGCCAGGTCACCGCCAAAGAAAGAAAGATGCTCTTAAGCC	2697
Db	7630	ACGATCACCTAATCCGCGCGCCAGGTCGCGCGCGGAAAGAAAGAAAGATGCTCTCCAAACC	7688
Oy	2698	AAAACTCTAGCTTAAAGAAAGAAAGAGCAACCCAAAGGAGGAAAGCAACGACCTTA	2755
Db	7690	AAAACTCTAGCTTAAAGAAAGAAAGAGCAACCCAAAGGAGGAAAGCAACGACCTTA	7748
Oy	2758	ACAGAGGAAACGACACGTAATGTATGTAAGTTGGAGTCGACAGACATTTCCGATAT	2817
Db	7750	ACAGAGGAAACGACAGCGTATGTATGTAAGTTGGAGTCGACAGACATTTCCGATAT	7809
Oy	2818	GCTGAACGGCCAAAGTAATGATATGCCCTGTGCGAGGAAGGCTGATGAACAACCT	2877
Db	7810	GTTGAACGGCCAAAGTAATGATATGCCCTGTGCGAGGAAGGCTGATGAACAACCT	7866
Oy	2878	CCAGCTTAAAGGAAATTTGATATATAGCAATTTACCGGCCGTGAATTTGAAGAGGCTAG	2933
Db	7870	CCAGCTTAAAGGAAATTTGATATATAGCAATTTACCGGCCGTGAATTTGAAGAGGCTAG	7928
Oy	2938	CATGTACGATTTGGAGTACGGCGACGTTCCCGCAGAACATGAATTCAGACACGCTCGATTA	2997
Db	7930	CATGTACGACCTTGGAGTATGGGAGCTTCCCGAATATGAATTCAGACACGCTCGATTA	7988
Oy	2998	CACCAAGCAAAACCAACCGGGCTTTCACACTGGCACAACGCGCGAGTCCAGATATGAGAA	3057
Db	7990	CACCAAGCAAAACCAACCGGGCTTTCACACTGGCACAACGCGCGAGTCCAGATATGAGAA	8048
Oy	3058	TGGGAGATTTACCGTACCGAGAGAGTGGCGGGAAGCCACACGCGAATCCGATCTCT	3117
Db	8050	TGGGAGATTTACCGTACCGAGAGAGTGGCGGGAAGCCACATGGAAAGCCGATCTCT	8108
Oy	3118	GGACAACAGAGCAGAGTTGTGTGCTATTGTTCTTAGAGGTGCAATATGAGGCAACGCGTAC	3177
Db	8110	GGACAACAGAGCAGAGTTGTGTGCTATTGTTCTTAGAGGTGCAATATGAGGCAACGCGTAC	8166
Oy	3178	GGCGCTTTCAGTGTCTACTTGGAAACCAAAAGGGGTGACATTTAGGATACCCCGGAAGG	3237

Db	8170	GGCGCTTTCAGTGGCTCACTTGGAAACAGAAAGGGGTGAACATCAAGATACC	8225
Qy	3238	TTCTGAACCGGTGCTCACTAGTTAAGCGCTATAGCGTGTCTCGAATGTCAGCTTCCCATG	3297
Db	8230	TTCTGAACCGGTGCTCACTAGTTAAGCGCTGTGGGTGCTTTCGAATGTCAGATTCCTCTTG	8289
Qy	3298	CGACAAACACCCCGTGTGCTATTCACTGACGCCAGAAACAACTCCGAGTGTCTGAGAA	3357
Db	8290	CGACAAACACCCCGTGTGCTATTCACTGACGCCAGAAACAACTCCGAGTGTCTGAGAA	8349
Qy	3358	GAAGCTGACAAATCCAAATTACGACAGCGTGTGGAGAAAGCTCTGAAATGTCATACAG	3417
Db	8350	GAAGCTGACAAATCCAAATTACGACAGCGTGTGGAGAAAGCTCTGAAATGTCATACAG	8409
Qy	3418	CCGGCCCCAAAGAAAGCATTAACCGATTCGATCACACTGACCAAGTCCCTCACTCGGGGTTCTG	3477
Db	8410	CCGGCCCCAAAGAAAGCATTAACCGATTCGATCACACTGACCAAGTCCCTCACTCGGGGTTCTG	8465
Qy	3478	CCCGTATTGCAGACACTCAACGCGCGTGTTCAGCCCAATATAAAATTGAGAACGTGTGGGA	3537
Db	8470	CCCGTATTGCAGACACTCAACGCGCGCATGTTTTAGCCCAATAAAATTGAGAACGTGTGGGA	8529
Qy	3538	CGAATCTGATATGATGATTCGATTAAGAAATTCAGGTCTCGGCAACAATTCGGCTCAATCAAGC	3597
Db	8530	CGAATCTGATATGATGATTCGATTAAGAAATTCAGGTCTCGGCAACAATTCGGCTCAATCAAGC	8589
Qy	3598	AGGCACCTGCGGATGTACCAAAATTCGCTTCATGTCTTCAGACCAAGCACTGACATCAA	3657
Db	8590	AGGCACCTGACAGACGTCAACCAAGTTCCGGTACATGTCTTACGACCAACGACATGACATCAA	8649
Qy	3658	GGAAAGACAGTATGAGAAAAATAGCTATACAGACACTTGGAACCTCGTGTCTTGGCCA	3717
Db	8650	GGAAAGACAGTATGAGAAAAATAGCTATTAAGTACATCTGGACCATGCGTGTCTTGGCCA	8709
Qy	3718	CAAAAGGATCTCTCTGTTAAGCTCAATGTCTCTCAGGTGACAGTGAACCGTCACTATATAC	3777
Db	8710	CAAAAGGATCTCTCTGTTAAGCTCAATGTCTCTCAGGTGACAGTGAACCGTCACTATATAC	8769
Qy	3778	GAGCGGAGCATCTGAGAAATTATCATCAACGTGTGAGAAAAAGATTCAGAGAAATTTGTGCG	3837
Db	8770	GAGCGGAGCATCTGAGAAATTATCATCAACCGTGTGAGAAAAAGATTCAGAGAAATTTGTGCG	8829
Qy	3838	TAGAGAGAGTACTGTGTTCCCAACCCGTCATAGAAAGCTGGTAAAGTCCACGTTTACGA	3897
Db	8830	TAGAGAGAGTACTGTGTTCCCAACCTGTCTCATAGAAAGCTGGTAAAGTCCACGTTTACGA	8889
Qy	3898	TCACTTTGAAGAGACGTCCTGCCGGTATCATTAACATGACAGGCCAGGCCACACGCGTA	3957
Db	8890	TCACCTTTGAAGAGACGTCCTGCCGGTATCATTAACATGACAGGCCAGGCCACACGCGTA	8949
Qy	3958	TAAGTCCATCTTGAGAGAAAGCGTACAGCGGAAGTGTATTAACAACCTTGTGGCAAGAA	4017
Db	8950	TAAGTCCATCTTGAGAGAAAGCGTACAGCGGAAGTGTATTAACAACCTTGTGGCAAGAA	9009
Qy	4018	CGTCACTTACGAATGTAAGTGTGCGCACTACAGACAGAGTATCTGTGACAGCGCAACGAA	4077
Db	9010	CGTCACTTACGAATGTAAGTGTGCGCACTACAGACAGAGTATGTAAGACAGCGCAACGAA	9069
Qy	4078	GATGAACGGCTGCCTAAAGCAAAACAGTGCATTGCTCTACAGAGCGCAACGAATG	4137
Db	9070	GATGAACGGCTGCCTAAAGCAAAACAGTGCATTGCTCTACAGAGCGCGCAACGAATG	9129
Qy	4138	GGTCTTCAACTGCGCGGATCTTATTAGGACACAGACCACTCAAGTGAAGGTAATGCA	4197
Db	9130	GGTCTTCAACTGCGCGGATCTTATTAGGACACAGACCACTCAAGTGAAGGTAATGCA	9189
Qy	4198	CATTTCATTCGCGTTGACACCGACAGTCTGCCGGTTCGTTTACCTCACAGCCTACAGT	4257
Db	9190	CATTTCATTCGCGTTGACACCGACAGTCTGCCGGTTCGTTTACCTCACAGCCTTACAGT	9249
Qy	4258	CACGAAGTGTTCAAAGGACATCACTCTTCACTGACTGCAATGCGAACCAATTTGCTGAC	4317
Db	9250	CACGAAGTGTTCAAAGGACATCACTCTTCACTGACTGCAATGCGAACCAATTTGCTGAC	9309

QY 4318 AACGAGAAATTTGGGGCTGCGAGCAGACGCAACGAGAAATGATTAAGGGTCTACATC 4377
| | | | |
Db 9310 AACGAGAAATTTGGGGCTGCGAGCAGACGCAACGAGAAATGATTAAGGGTCTACATC 9369
QY 4378 CAGGAATTTTCTGTGGGGCGAGAGGGCTGAGTACGTATGGGGTAAACATGAACGAGT 4437
| | | | |
Db 9370 CAGGAATTTTCTGTGGGGCGAGAGGGCTGAGTACGTATGGGGTAAACATGAACGAGT 9429
QY 4438 CAGAGTCTGGGGCCAGAGAGTGGGACCAAGGAGACCCCATGAGATGGCCGATGAGATCAT 4497
| | | | |
Db 9430 CAGAGTCTGGGGCCAGAGAGTGGGACCAAGGAGACCCCATGAGATGGCCGATGAGATCAT 9489
QY 4498 CATCCATATATATATGAGGAGTCCAGTCTACACTGTATGCTGTGTGTGCTGCTCT 4557
| | | | |
Db 9490 CATCCATATATATATGAGGAGTCCAGTCTACACTGTATGCTGTGTGTGCTGCTCT 9549
QY 4558 TGCTATCTGGTAGCACTGATCATGACAGCTTGATCGCCCAAGCAAGAGAGAGATG 4617
| | | | |
Db 9550 TGCTATCTGGTAGCACTGATCATGACAGCTTGATCGCCCAAGCAAGAGAGAGATG 9609
QY 4618 CTTGACGCCATACCGCGCTTGACCGGAACGATACCCAGACATTAAGGGTGTG 4677
| | | | |
Db 9610 CTTGACGCCATACCGCGCTTGACCGGAACGATACCCAGACATTAAGGGTGTG 9669
QY 4678 CTGCAATCGGCGCAACCAACGCTGAAACATTTGAGAACTTTGAAACCATCTGTGTTTAA 4737
| | | | |
Db 9670 CTGCAATCGGCGCAACCAACGCTGAAACATTTGAGAACTTTGAAACCATCTGTGTTTAA 9729
QY 4738 CAACCAACCGTTTCTGTGGGACAGTTGTGATTCCTGTGGACGCTTGTATTTCTGT 4797
| | | | |
Db 9730 CAACCAACCGTTTCTGTGGGACAGTTGTGATTCCTGTGGACGCTTGTATTTCTGT 9789
QY 4798 CCGCTGCTTTTCAATGCTGATGCTTTTATTTGTTGTTGAGGGCTGTGCTGGGAAAGT 4857
| | | | |
Db 9790 CCGCTGCTTTTCAATGCTGATGCTTTTATTTGTTGTTGAGGGCTGTGCTGGGAAAGT 9849
QY 4858 AGACGCTTGCAGAAATGCGAGCACTGTCGAAATGTTCCGGGAGTCCGTTAAAGGCTT 4917
| | | | |
Db 9850 AGACGCTTGCAGAAATGCGAGCACTGTCGAAATGTTCCGGGAGTCCGTTAAAGGCTT 9909
QY 4918 GGTGGAACGCGCAGGTTACCGGCACTTAACTGGAATCAAGGTCGTCTTCAGGAAAT 4977
| | | | |
Db 9910 GGTGGAACGCGCAGGTTACCGGCACTTAACTGGAATCAAGGTCGTCTTCAGGAAAT 9969
QY 4978 AACCACTTCACTAACAAGAGTACGTGACCTGCAATTTCAACAGTCAATTCCTTCAAC 5037
| | | | |
Db 9970 AACCACTTCACTAACAAGAGTACGTGACCTGCAATTTCAACAGTCAATTCCTTCAAC 10029
QY 5038 ACAAGTTAAATGCTGCGGGTCCCTCGAGTGCAGGCAATCCCAAGGCGATTAACATG 5097
| | | | |
Db 10030 ACAAGTTAAATGCTGCGGGTCCCTCGAGTGCAGGCAATCCCAAGGCGATTAACATG 10089
QY 5098 CCGCGTTTTTGGCGGTGTACCTTTTCAATGTTGGGAGGCGCACATGCTTCTGTGACG 5157
| | | | |
Db 10090 CCGCGTTTTTGGCGGTGTACCTTTTCAATGTTGGGAGGCGCACATGCTTCTGTGACG 10149
QY 5158 TGAAGAACACACACTGAGTGAAGGCGTACGTTGCTTCAAGTGCATTAATGATCA 5217
| | | | |
Db 10150 TGAAGAACACACACTGAGTGAAGGCGTACGTTGCTTCAAGTGCATTAATGATCA 10209
QY 5218 CGCAGTGCACATAAAGTTACACAGGCTCTGAAAGTGGGCTGGGCTATAGTATACG 5277
| | | | |
Db 10210 TGCAGTGCACATAAAGTTACACAGGCTCTGAAAGTGGGCTGGGCTATAGTATACG 10269
QY 5278 CAACACCAACCGGACCTGATAGCTTTGTACATGCGCTCAACGCAAGGTTCTTCAACGGA 5337
| | | | |
Db 10270 CAATACCAACCGGACCTGATAGCTTTGTACATGCGCTCAACGCAAGGTTCTTCAACGGA 10329
QY 5338 CTTGAAAGTCAATAGCAGGCGGATATACGCGCTTTTCAACCTTTGACCAATAGTGTCT 5397
| | | | |
Db 10330 CTTGAAAGTCAATAGCAGGCGGATATACGCGCTTTTCAACCTTTGACCAATAGTGTCT 10389

QY 5398 CATCAGAAAGGGGCTTTGTTACACTACGACTTCCCTGAGTATGAGCTATGAACACAG 5457
| | | | |
Db 10390 CATTAAGAAAGGGGCTTTGTTACACTACGACTTCCCTGAGTATGAGCTATGAACACAG 10449
QY 5458 AGGTTTGGCGGATTTTCAAGCAATCTCGCTGTTATGCTACAGACATAGTAGCCGCACTGA 5517
| | | | |
Db 10450 AGGTTTGGCGGATTTTCAAGCAATCTCGCTGTTATGCTACAGACATAGTAGCCGCACTGA 10509
QY 5518 CATACGCGCTGTAAGCTTCTGTCAAGAACATCCAGTCCCTTCAACCCCAAGCAGATATC 5577
| | | | |
Db 10510 CATACGCGCTGTAAGCTTCTGTCAAGAACATCCAGTCCCTTCAACCCCAAGCAGATATC 10569
QY 5578 AGGATATGAATGTGAAGAAACAATCAGAGCAACCCCTGCAAGAAACAGACCATTTGG 5637
| | | | |
Db 10570 AGGATATGAATGTGAAGAAACAATCAGAGCAACCCCTGCAAGAAACAGACCATTTGG 10629
QY 5638 ATGTAAATTTGAAGTGAAGCTTCTGAGAGCTTAACTGTGCTTACGGGCAATCCCTAT 5697
| | | | |
Db 10630 ATGTAAATTTGAAGTGAAGCTTCTGAGAGCTTAACTGTGCTTACGGGCAATCCCTAT 10689
QY 5698 CTGCAATGACATCCCTGATGACGCTTTTGTGATGATCATGAAATCAACCAATTTTGA 5757
| | | | |
Db 10690 CTGCAATGACATCCCTGATGACGCTTTTGTGATGATCATGAAATCAACCAATTTTGA 10749
QY 5758 AGTTAGCTGACAGTAGCAGACTGCAATTTATCTGAGACTTTTGTGTCTTAACAT 5817
| | | | |
Db 10750 AGTTAGCTGACAGTAGCAGACTGCAATTTATCTGAGACTTTTGTGTCTTAACAT 10809
QY 5818 ACAATCAAAAGCTGACAGGAGGAGACATTTTCACTTCACTCCACGACGCTGT 5877
| | | | |
Db 10810 ACAATCAAAAGCTGACAGGAGGAGACATTTTCACTTCACTCCACGACGCTGT 10869
QY 5878 TTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5937
| | | | |
Db 10870 TTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10926
QY 5938 ATGAGCCCAAGCAAGCAAAATTTTATGTTTGTGATGCGGCAAGAGTCCACTGATGC 5997
| | | | |
Db 10927 ATGAGCCCAAGCAAGCAAAATTTTATGTTTGTGATGCGGCAAGAGTCCACTGATGC 10986
QY 5998 TGAATGTAACCAACCGGCGGACCAATTAATGAGAAACCAATTAAGTCAACCAAGAT 6057
| | | | |
Db 10987 TGAATGTAACCAACCGGCGGACCAATTAATGAGAAACCAATTAAGTCAACCAAGAT 11046
QY 6058 CCAAGCGGCGGAGTTTCCAAACATCTTGGAGCTGGCTGCTGCACTGTTTGGGGAGCATC 6117
| | | | |
Db 11047 CCAAGCGGCGGAGTTTCCAAACATCTTGGAGCTGGCTGCTGCACTGTTTGGGGAGCATC 11106
QY 6118 ATCCCTCATTTGTTAGAGACTTAATAGTGTGTCTGACGCTTATGCTTAACACAG 6177
| | | | |
Db 11107 ATCCCTCATTTGTTAGAGACTTAATAGTGTGTCTGACGCTTATGCTTAACACAG 11166
QY 6178 TAGATGCTAGCGGCGGACACTGACATAGCGGT-AAAATGCTGATGTTCTTCCAGGAAAC 6236
| | | | |
Db 11167 TAGATGCTAGCGGCGGACACTGACATAGCGGTAAAATGCTGATGTTCTTCCAGGAAAC 11226
QY 6237 GTGTGCTATTAATGCCACGCGCGCTTGACACTAAATCTGATGATTTCCGAGGAAAGCAC 6296
| | | | |
Db 11227 GTGTGCTATTAATGCCACGCGCGCTTGACACTAAATCTGATGATTTCCGAGGAAAGCAC 11286
QY 6297 AGTGCATTAATGCTGTGAGAGTGCATTAATGCTATTAATGCTATTAATTAACACACT 6356
| | | | |
Db 11287 AGTGCATTAATGCTGTGAGAGTGCATTAATGCTATTAATGCTATTAATTAACACACT 11346
QY 6357 ATATCACTTTATGAGACTCACTATGAGTCTTAATTAATTAATTAATTAATTAATTAATTA 6416
| | | | |
Db 11347 ATATCACTTTATGAGACTCACTATGAGTCTTAATTAATTAATTAATTAATTAATTAATTA 11406
QY 6417 AAACACTATACACACTTAATTAATTTCTTTTATTAATTTTATTTATTTTATTTTATTT 6476
| | | | |
Db 11407 AAACACTATACACACTTAATTAATTTCTTTTATTAATTTTCTTTTATTTATTTTATTT 11466
QY 6477 TAAATTTT 6484

Db 11467 TTAATTT 11474

RESULT 2

US-08-801-263A-7

Sequence 7, Application US/08801263A

Patent No. 5811407

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Simpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

NUMBER OF SEQUENCES: 12 Expression of Heterologous Genes in the Bone Marrow

CORRESPONDENCE ADDRESS:

ADDRESS: Bell Seitzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 5811407th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,263A

FILING DATE: 19-FEB-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 11663 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-801-263A-7

Query Match 37.0%; Score 2397.6; DB 1; Length 11663;

Best Local Similarity 64.6%; Pred. No. 0;

Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;

QY 580 ACTGACGATGAGACCGGAGCGTATATTTCTCATCGGAAACAGGCCAGGTCACTTC 639

Db 5707 ACTGCTAACCGGGGTAGGTGGTACATATTTTCACGAGACAGGCCCTGGGCACTTGC 5766

QY 640 AACAGAAATCAAGTCAATGTAATCAAGAACTATATGATGGGCGGCTCCATG 699

Db 5767 AAAAAGATCGCTTGTGCAAAACAGCTTACAGAACCGACCTTGAGGCGCAATGTTCTGG 5826

QY 700 AGAATATTTAGCCCCCGGCGCTCGATCTCGAAGAGAGAAATGTTTACAGAAAGTGC 759

Db 5827 AAAAGATCTAGCGCCCGGTGCTCGACGTCGAAAGAGAGACAGCTCAAACTCAGATACC 5886

QY 760 AATTATGCGCTCTTGAAGAAATGAGAGAGATATCAACAGAAAGTAGAAAAATATGA 819

Db 5887 AGATATGCGCCCAACCGAACCAAAAGAGAGATACAGTCTGAAAGTAGAAAAACCGA 5946

QY 820 AAGCAATTAACGAGGAGGAGCTATTTCTGATTTGGGCAATATCTATCATAGAGTGA 879

Db 5947 AAGCATTAACGAGGAGGAGCTATTTCTGATTTGGGCAATATCTCTACAGTGAAGTGA 6003

QY 880 ATCTGTGAGTGTATACAGAGTCAATTTATCTGTACCAATCTACTCTGTCACGTTAATTA 939

Db 6004 ATCAAGCAGAAATGCTATAGATCACTTACCGGAAACCATGATTCACAGGTATACAG 6063

QY 940 AAGGTTTAACTCTGACAGGTCGGGTTAAAGCTGCACTTATTCAGAGAAAT 999

Db 6064 CGAATCTTGAACCCAAAGTTTGTGAGTCTGTTTGAACCACTATCTGCAATGAGAAAT 6123

QY 1000 ACCCTACAGTACCGGATTTGATTAACAGATGATGATGATGATGATGATGATGATGATG 1059

Db 6124 ACCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6183

QY 1060 ACGGCGATGCTGCTGTCTGATATGATGATGATGATGATGATGATGATGATGATGATG 1119

Db 6184 ACGGACAGTGGCTGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 6243

QY 1120 CAAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1179

Db 6244 CGAAAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6303

QY 1180 ATACATTAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239

Db 6304 ACAGCTTCAAAACAGTGTCTCATTTGCGGACTTAAAGAAATGCAACCTCACAAATG 6363

QY 1240 GAGAAATTAAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1299

Db 6364 GTGAATCTCAAAACAGTGTCTCATTTGCGGACTTAAAGAAATGCAACCTCACAAATG 6423

QY 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359

Db 6424 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6483

QY 1360 TTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419

Db 6484 TTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6543

QY 1420 ATATCTTAAACCGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1479

Db 6544 ATATCTTAAACCGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 6603

QY 1480 ATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1539

Db 6604 ATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6663

QY 1540 AGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599

Db 6664 AGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6723

QY 1600 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1659

Db 6724 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6783

QY 1660 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1719

Db 6784 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6843

QY 1720 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779

Db 6844 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6903

QY 1780 ACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1839

Db 6904 ACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6963

QY 1840 CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1899

Db 6964 CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7023

QY 1900 TCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1959

Db 7024 TCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7083

QY 1960 AACGATTAACAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAG 2019

Db 7084 AACGATTAACAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAG 7143

OY	2020	TCGTCCTCGGAACCTTGATGTCGCGGAGAAATGCGCACCTTGGCTGTGAACTGGAAATGAAAA	20739
Db	7144	TAGTATCTGACAAGAAATGCGTGAGAGGTGGCCACTGGCTCAACATGAGAGGTTAAAGA	72030
OY	2080	TTATTTGATGACAGTTATTTGGTATCAAGACACCTTACTGTGTGGGGATTTATTCCTGTGTHG	21339
Db	7204	TCATTGACGCAAGTATATCGGCGGAGAACACACCTTACTTCTGGGGGTGGATTACTTCTTGCAAG	72633
OY	2140	ACCAAGTAAACAGGCAACACCTGCAAGATCGCAGAACCTCTTAATAAAGGCTTTTAAAGCTTG	21399
Db	7264	ATTGCGTTACTTCCACACAGCTGTGTCCGGTGGGAGCCCTTGAATAAAGGTGTGTTTAAAGTTGG	73233
OY	2200	GAAAAACATTTGCGCAGTGCATGATATACCAAGACTGCGACCCGCGCCGGGCACTGCATGATGT	22539
Db	7324	GTTAAACCCCTCCGACGCCACGATGAGCAAGAGAAAGAACAGAAAGACGCGCTCTCTCAAGATG	73883
OY	2260	AAGCAATGCGATGGAACAGAAATTTGGAATTCGAGACAGTTAGTAGAAGGCGGTAGAAATCCA	23139
Db	7384	AAACAAAGGCGGTGTTTAGTAGAGGTATTAACAGACACTTAGACAGTGTGCGGTGGCAATCTC	74433
OY	2320	GATACGATCAATACTGCGACGCGCATCATCAAGTCTCTGTGACAGTTACCGGAAAGCG	23739
Db	7444	GGTATGAGGTGAGCAACATCACACTGTCTGTGCTGCGCATTTGAGAACTTTTGGCCAGACACA	75030
OY	2380	TTAAGAACTTCAAGACATTAAGAGGAGGCCCAATCACCTCTTACCGGCTGACCTTAATATAG	24339
Db	7504	AAAGACATTTCAAGCCATCAGAGGGGAAATTAAGCATCTCTAGGTGTCTTAAATATGT	75633
OY	2440	TGACCTAGTAGA-----CAGCACCTAACCCACCGGAGAAATTTTCCA	24882
Db	7564	CAGCATATTAACATTTTCATCTGATCTAATAACAAACACACACACATGAAATGAGAGATTTCT	76233
OY	2483	TACCTCTAGCTGAACCTTTCACACAGTTTAACTTACCAATTCGAGATGGCTTACCGAGATCCA	25424
Db	7624	TTAATCATGCTGGCGCGCGCCCTTCCAGCCCCCACTGCGCATGTGAGAGGCGCGGAGAA	76883
OY	2543	AACTCTCTAGGCGCGCTGAGAGGCGCTTTGGCGCCCGCTGTGCTCAATATGAGAGAT	26020
Db	7684	GGAGGCAAGCGGCGCCCGATGCTGCTGCCCAATGGCGTGGCTTCCCAATTCAGCAACTGA	77433
OY	2603	CTTAGGAGGTGGATGTGTAACCTGTACCTTTCAAAACAAC--GATCACTTAATCGCGCGCAG	26666
Db	7744	CCACAGCCGTGATGCTCTTACTGATGAGCAAGCACTAGACCTTAAACCCCAAGCCCTAC	78030
OY	2661	GTCCACGCGCAAGAAAGAAAGAAAGTGTCTTAAGCCAAAACCTTACTGAGCTTAAAAAGA	27202
Db	7804	GCCCCCGCCCGCGCCAGAAAGACAGGCGCCAAAGCAACACCGAAGCGGAAAGAAACCA	78633
OY	2721	AGAGCAACCAAGCCCAAGAGACGAAAC---GCAGACTTAACCGGGGAAAGACAAACCTA	27777
Db	7864	AAACACAAGAGAGAAAGAAAGAACCACTGCAAAAACCAAAACCCGGAAGAGACAGCTTA	79233
OY	2778	TGTGTATGAAGTTGGAGTCGCAACAAGACTTTCCGATC---ATGTGTAAGCGGCCAAAGTGA	28343
Db	7924	TGGCACTTAAAGTTGGAGGCCACAGACCTGTTGCAAGTCAAAAATAGAGGACCGAAGATGTCA	79883
OY	2835	ATGATATATGCTCGCTGTCTCGAGAGAGCTGATGAAAACACTCCACTCACTTTAAGGAAAAA	28944
Db	7984	TGCGGCAACGCACTGCGCAATGGAAGGAAAGTATATAAACCACTCCACTGTAAGGAACTA	80433
OY	2895	TTGATTAATGAGCAATTAGCGGCGCGGAATTGAAGAAAGGCTAGACATGTACACTTTGGAGT	29544
Db	8044	TTGACCAACCTGTGTATCTAAAGCTCAAAATTCACCAAGTCTGTACGATACATGAGATGAGT	81030
OY	2955	ACGGGCAAGTTCCCGAAGACATGAATCAGACACGCTCAGTACACACGACAAACCAAC	30143
Db	8104	TCGCAACAAGTTCGGGTCAACATGAGAAAGAGGCGTTCACTTACACGATGAACACCTGT	81633
OY	3015	CGGGCTTTCAACATCTGGCAACACGCGCGCATCTCAATATGAGAAATGGAGATTTTACCTGATC	30744
Db	8164	AAGGTTTTCACATCTGGCAACACGAGGCGGTGCAGTATATGTGAGAGCGAGATTTTACATCC	82233

QY	3075	GGAGAGAGTGGGGGGGAAAGGCGACAGCGGAAACCGATCTCTGGACAAACAGAGCGAG	3134
Db	8224	CCCGGGGAGTGAAGGCGAGAGAGACAGTGGTCTCGGATTATGTAATCTCAGGCGCGG	8283
QY	3135	TTGTGGCTATTGTTCTTAGAGGTGCAATATGAGGCGACGCGTACGCGCTTTCAGTGGTCA	3194
Db	8284	TTGTGGCATAGTCTCTCGAGGGGGCTATGAGGGAAACAAGAACCGCTTTCGGTCTGTA	8343
QY	3195	CTTGGAACCAAAAGGGGTGACCATTAAGCATCCCCGAAAGTTCTGAAACGTTGT---	3251
Db	8344	CCTGGAAATGCAAAAGGGGAAAGCAATCAAGACAAACCCGGAAAGGAGCAAGAAAGTGTCTG	8403
QY	3252	-----CACTAGTTTACAGCGGCTTAAGCGGCTTGGAAATGTCAGCTTCCCATGGCAAAAC	3305
Db	8404	CTGCACCATGCTGTACCGGCATGTGCTTGTGGAAACGTAGCTTCCCATGTACATGCC	8463
QY	3306	CACCGGTGTGCTATTTCATCTGACGCGCAGAACCAACACTGACGTGCTCGAAGAGAACGTGC	3365
Db	8464	CGCCCACTGTTCACCCCGGACCAATCCAGAGCTCTGCACATCTCGAAGAGAACGTGA	8523
QY	3366	ACAATCCAAATTACGACACGCTGTGTGAGAAAGTCTTGAAT--GTTCATACGCGCGC	3422
Db	8524	ACCAAGAGGCTTACGACACCTGTCTCAACGCATATTGCGGTGGGATGTCGCGGCAAGAA	8583
QY	3423	CCAAACGAGATTACCGATATCTTCACTGACAGCACTCCCTACTCTGGGTTCTGCCGT	3482
Db	8584	GTAAGAAGAGGTCACTGACACATTTACTTTGACACCGCTACTCTGGGCAATGTCGT	8643
QY	3483	ATTGCAGACTCTCAACGCGCTGTTCAGCCCAATPAAATTTGAGAACGTGGGAGCAAT	3542
Db	8644	ACTGTACCACTACTGAAACGTTGCTTTAGCCGATTTAAGATGAGCAGGTCTGGGATGAG	8703
QY	3543	CTGATGATGATCGATTAGATCCAGGTCTTCGAGCAATTCGGCTACATTCAGGCAAGCA	3602
Db	8704	CGAGACGAACAACCACTACGATACAGACTTCCGCCAGTTTGATAGCAGCAAGCGGAG	8763
QY	3603	CTGCGGATGTCAACCAATTCGTTACATGTTTTCGACACAGACATGACATCAAGAGAG	3662
Db	8764	CAGCAAGCTCAAAATAGTACCGGTCACTATGTGCTCGAGCAGAGATCAATCTGTCAAAAGG	8823
QY	3663	ACAGATGAGAAATAATAGTCTATCGACATCTGAGACCCGTGCGTCTTGGCCACAAG	3722
Db	8824	GCACCATGATGACATCAAGATCGACACTGAGACCGTGTGAAGGCTTAGTACTCAAG	8883
QY	3723	GGTACTTCTGTAGCTCAATGTCCTCCAGGTGACAGTGAACCGTCAGTATCCAGCG	3782
Db	8884	GATACTTCTCTCGCGAAAGTGTCTCCAGGGGACAGGTTAACGTTAGCATCGAGTA	8943
QY	3783	GAGCATCTGAATAATTCATGACACCGTGGAGAAAAAGATCAGAGAGAAAGTTGTCCGTAGAG	3842
Db	8944	GCAACTCAGCAACGTCAAGCAATATGCGCCGCAAGATPAAACCAAAATTCGTGGGAAGCG	9003
QY	3843	AGGAGTATTGTCCACCGGTCCATGGAGAAAGTGTAAGTGCACGTTTACGATCACT	3902
Db	9004	AAAAATATGACCTTACTCTCCGTTACCGTTAAGAAAGATTCTTTCGACAGTGTACACCGTC	9063
QY	3903	TGAAGGAGACCTGCGGGGTACATTAACATGACAGGCGCAGGCCCAACGCGTATAAGT	3962
Db	9064	TGAAGAGAAACAACCGCGGTACTACACTATGACACAGGCGGGAACCGCACTATACAT	9123
QY	3963	CCTATCTGGAGAAAGCGTCAGGCGAAGTGTACATTAACAACCTTCTGGCAGAAAGCTCA	4022
Db	9124	CCTATCTGGAGAAATCATCAGGAGAAAGTTTACGGAGGCCACATCCGGGAGAAACATTA	9183
QY	4023	CTTACGAATGTAAAGTGTGGCGACTACAGACACAGTATCTGTGAGACACCGCAACGAATGA	4082
Db	9184	CGTACGAATGTCAAGTGGCGGCAATTACAGACCGGAACCGTTTACGACCCGTACCGAAATCA	9243
QY	4083	ACGCGTGCATTAAGCAAAACAGTGCATGTGCTTCAAGAGCGACCAACGAATAGGCTCT	4142
Db	9244	CGGGGTGACCGCCATCAAGCAGATGTGCTGCTTAAAGAGCAACCAACGATGGGCTCT	9303
QY	4143	TCAACTCCCGGATCTTATTAGGCAACAGACCACTCAGTGCACAGGTAAATGTGCATTC	4202

Db 9304 TCAACTGCGCCGACCTGATCGAGACACGCCGACACACGCCCAAGGAAATTGCAATTGGC 9363
Qy 4203 CATTCGCGTTGACACCGACAGCTGCGCCGGTTCCGTTAGTTCACACGCTTCACAGCACA 4262
Db 9364 CTTTCAAGCTATCCCGAGTACCTGACATGATCCCTGTTGCGACGCGCGAAGCTAGTAC 9423
Qy 4263 AGTGGTTCAAAGGACATCACTCCACCTGACCTGACATGACGACCAATGCTGACACGA 4322
Db 9424 ACGGCTTTAAACATCATGAGCTCCAAATTAGACACAGACCATCTGACATTGCTTCACCA 9483
Qy 4323 GAAATTTGGGCTGCGACAGACGACACAGAAATGATTTACAGGTTCTACATCCAGA 4382
Db 9484 GGAACATGAGGGGCAAAACCCGACCAACCACTGAATGATCATCGGAAACAGGTTGAA 9543
Qy 4383 ATTTTTCGTGGGGGAGAAAGGCTGAGTACGTATGAGGGGTAAACATTAACCATGACAG 4442
Db 9544 ACTTCACCGTGCACCGAATGAGCTCGAATATATAGGGGCAATCAGAACCATGAAGGG 9603
Qy 4443 TCTGGGCCAGAGATCGGACACGAGCGACCATGATGATGAGCCGATGATCATCTCC 4502
Db 9604 TCTATGCCAAGATCTGACACAGAGACCTTCACGATGAGCCACACGAATATGATACGC 9663
Qy 4503 ACTATATCATCGCATCCAGCTTACACTGTCAITGTGCTGTGTGTGCTCTTGTCTA 4562
Db 9664 ATTACTATCATCGCATCTCTGTATACCATCTTAAGCCGTGCATCAGCTGCTGTGGCA 9723
Qy 4563 TCCGTGATGAGCATGATCATACAGCTTCATCGCAAGCAAGCAAGAAACATGCTTGA 4622
Db 9724 TGAATATGGGTAACTGTGTGACGATATGTGCTGTAAAGCGCGGTAGTGGCTTGA 9783
Qy 4623 CGCCATACGCGCTTGCACACGACGCAAGGTAACCCACAGCATTTAGCGTTTGTGCTCA 4682
Db 9784 CGCCATATGCGCTTGGCCCAATGCGGTGATTCACACTTCGCTGCACTTTTGTGCTGTG 9843
Qy 4683 TTGCGCCAAACCAACGCTGAAACATTTGAGAAACCTTTGAAACATCTGTGTTTAAACA 4742
Db 9844 TTAGTGGGCTTAAATGCTGAAACATTCACCGAGACATGATGTTATGTATGCAACAGC 9903
Qy 4743 AACGTTTCTGTGGGACAGTTGTGATCTCTGTGGAGCGCTGTATTTCTGTTCCGCT 4802
Db 9904 AGCGTTCTTCTGTGGGTCAAGCTGTGTATACCTGTGGCGCTGTGCTTTTAAATGCGCT 9963
Qy 4803 GCTTTTATGCTGATGCTCTTTTATATGTTGATGAGCGCTGTGCTGTGGGAAAGTAAAG 4862
Db 9964 GTTGTCTATGCTGCTGCTTTTATGTTGATGAGCGCTGTGCTGTGCGAAAGTAAAG 10023
Qy 4863 CTTTTCGAAATGCAACACTGTGCAAAATGTTTGGGGATCCCGTAAAGGCTTGTGCTG 4922
Db 10024 CCTACGAACATGCGACACTGTTCMAATGTGCCACAGATACCGTAAAGGCACTTGTG 10083
Qy 4923 AACGCGAGGTTATGCGGCACTTAACTGAGATACAGGTGTCTCATCGAATTAACAC 4982
Db 10084 AAAGGAGAGGGGTAGCCCCGCTCAATTTGAGATTAATCTGTATGCTCTCGAGGTTTGC 10143
Qy 4983 CTTCAACTTAACAAGAGTAGTGAAGCTGCAATTCACAGATCATCTTCAACACAG 5042
Db 10144 CTTTCAACCAACAGAGATGATTAATCTGCAATTAACACAGTGTGCTCTTCCCTTAAG 10203
Qy 5043 TTAATGCTGGGCTCTCTGAGTGAAGGATCTCAAGGCGGATTAACATGCGCG 5102
Db 10204 TCAATGCTGGGCTCTCTGAGATGAGCCGCGCTCAAGAGATTAATCTGCAAGG 10263
Qy 5103 TTTTTCGCGGTGTATCTTTTCAATGTGGGAGCGCAATGCTTCTGTGACAGTGA 5162
Db 10264 TCTTTGAGGGGTATACCTCTTCAATGTGGGAGGAGCAAAATGTTTTCACAGTGA 10323
Qy 5163 ACACACAATGAGTGAAGGCTAGCTGAGTTGCTCCAGCTGACCTTAATAGACAGAG 5222
Db 10324 ACAGCAGATGATGAGGCGTATGCTGAAATGTCTAGTATGAGTGGCGCACTGACGCGC 10383
Qy 5223 TCGCACTAAAATTCACACAGCTGTCTGAAAGTGGCGCTGATATAGTATACGCAACA 5282

Db 10384 AGCGATTAAGTGTACATCTCCCGATGAAGTAGAGATGCGTATATGTATCGGGAACA 10443
Qy 5283 CCACCGGACCTCGATATCGTTTGTCAATGGCGTCAAGCGAGTTTCCACAGGACCTGA 5342
Db 10444 CTACAGTTTCTATGATGTGTATCGTAAACGAGTCAACAGGAACGTTTAAAGCTTGA 10503
Qy 5343 AGGTATAGCAGGCGCATATACCGCTTGTACACCTTGTGACCATTAAGTGTCTATCA 5402
Db 10504 AAGTATAGCTGAGCAATTCATGACATTTGTTTACACATTTGATCAAGAGTCTGTATCA 10563
Qy 5403 GAAAGGCGCTTGTATACATACGACTTCCCTGATATGAGTATGAACCAAGAGCGT 5462
Db 10564 ATCCGCGCTGTGTACAACTATGATCTTCCGGAATCGAGCGATGAACAGAGCGT 10623
Qy 5463 TCGGCAATTAACAAGATCCCGCTGATGTATAGATGATAGTATGCGCGCATACATAC 5522
Db 10624 TTGAGACATTCAGACTACCTCTTGACTGACAAAGACCTATGCGCAGACAGCATTA 10683
Qy 5523 GCGTGTGAAGCTTCTGTCAAGAACTCAACGTCCCTACACCCAAAGCAATACAGGT 5582
Db 10684 GGTACTCAAGCTTCCGCAAGACGTGATGCTGATACGAGCGCGCATCTGTGAT 10743
Qy 5583 ATGAATGTGAGAGACACTCAGAGAGACCCCTGCAAGAAACAGACCAATTTGATGA 5642
Db 10744 TCGAGATGTGAAAAAACAACACTCAGCGCGCCACTGACAGAAACCGCCCTTTGGTGA 10803
Qy 5643 AAATTTGAAGTGAACCTGTGAGCGTCTAATGTGCTTACGGGCAATCCCTATCTCA 5702
Db 10804 AGATTTGATCAATCCCTGTGAGCGGTGATGCTGATACAGGAAACATTCCTCATTTCTA 10863
Qy 5703 TTGACATCCCTGATGACAGCTTTTGTGATATCATGATCATCAACATTTTAAAGTTA 5762
Db 10864 TTGACATCCCGAAGCTGCTTTTATCAGACATGATGACACCACTGCTCAACAGTCA 10923
Qy 5763 GCTGACAGTAGCAGACTGATTTATCTGAGACTTGTGTGTTCTTAACTTAACAT 5822
Db 10924 AATGTATGTACAGTAGTGCATTAATCAGGAGACTTCGAGGATGAGCTACCTGCA 10983
Qy 5823 ACAAGCTGACAGGAGGACATTTGTCAAGTCACTCCACTCCACAGCAGCTGTTTGA 5882
Db 10984 ATGTATCGAGCGGAGAGCAATGCTGTGACATTTGCAATTCAGACACAAACCTCC 11043
Qy 5883 AGAAGGACACACATGATGCTGCTGAGCAGACATCACTATTTTAGCATGCA 5942
Db 11044 AAGTGTGACAGTTCAATGCTGTGAGAAAGAGGAGTGAAGTCACTTACGACCGCA 11103
Qy 5943 GCCCACAAGCAATTTATGATTTGCTATGCGGCAAGAACTCACTGCAATGCTGAT 6002
Db 11104 GCCCAGGCGAACTTATGATGCTGTGTGTGTAAGAGACAAATGCAATGCAAGAT 11163
Qy 6003 GTAAACCAAGCGGCGACCACTAATTTGAGAACCAATTAAGTGCACCAAGATTTCCAG 6062
Db 11164 GCAACCAACAGCTGATCATATGCTGAGCAACCCGCACAAAATGACCAAAATTTCCAG 11223
Qy 6063 CGGAGTTTCAAAACATTTGGAATGGCTGTTGCACTGTTTGGGGAGCATCATCC 6122
Db 11224 CCGCATCTCAAAACATTTGATGAGTGTGCTGTTTGGCGGCGCTGTGCTG 11283
Qy 6123 TCAATGTTGAGAGCTTAATGATGTTGTGCTGACCTATATGCTTAACACAGATGAT 6182
Db 11284 TATTAATTAATGAGCTTATGATTTTGTGCTGACAGATGCTGATGACAGAGAT 11343
Qy 6183 GACTGAGCGGAGACATGTA-CATAGCGGTAAACTGATGATCTTCCGAGGAAGCGTGT 6241
Db 11344 GACCGTACGCGCCCAAGACCCGACAGCAAAATCTGATGATCTTCCGAGGAAGTGT 11403
Qy 6242 GCATATGCGACGCGCGCT-----TGACACT 6268
Db 11404 GCATATGCTACAGCTGTGTATATTAATCCCGCTTACCGCGGCAATATAGCAACACC 11463
Qy 6269 AAAATCTGATGATTTTCCAGGAAGACAGTGCATTAATGCTGTGCACTGT 6318
Db 11464 AAAATCTGATGATTTTCCAGGAAGCGCAAGTGCATTAATGCTGTGCACTGT 11513

RESULT 3

US-09-102-248-7

Sequence 7, Application US/09102248
Patent No. 6008035

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Simpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell Seltzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 6008035ch Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/102,248

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/801,263

FILING DATE: 19-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 11663 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-102-248-7

Query Match 37.0%; Score 2397.6; DB 3; Length 11663;

Best local Similarity 64.6%; Pred. No. 0;

Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;

QY 580 ACTGACGGTATGAAACGGGAGCGTATATTTTCATCGAAACAGGCCAAAGTCACTTC 639
DB 5707 ACTGTCMAACGGGGGTGAGTGGTACATATTTTCAGACGACACGCGCTGGGCACTTGC 5766
QY 640 AAGAGAAATCAGTACGTCATGTAATACTACAGAAACCTATATTGGAGCGGCGCTCATG 639
DB 5767 AAAAGAAAGTCGTTCTGCAAAACCAAGCTTACAGAAACGACCTTGGACGGAATGTTCTGG 5826
QY 700 AGAAGTATACGCGCCGCGCTCATCTCGAAAGAGAGAAATGTTACAGAAAGAACTGC 759
DB 5827 AAGAAATCTACGCGCGCGCTGCTGACACGTCGAAAGAGAAAGCACTCAAACTCAGATAC 5886
QY 760 AATTATGCGCTTGAAGAAATGAAGAGATATCAATCAAGAAAGTAAATATGA 819
DB 5887 AGATGATGCGCCAGCAAGCAACAAAGCAGTACCACTCGAAAAGTAAAGAAACAGA 5946
QY 820 AAGCAATTAAGCGGAGCGCATCTTTCTGATTTGGGCAATATCTATCAATCAGAAATGA 879
DB 5947 AAGCCATTAACCACTGAGCGCATCTTTCTGAGGGCTACAGGCTGTAATCACTGCCACAG-- 6003
QY 880 ATCTGTGAGTGTACAGATCAATATCTCTGTAACAATCTACTCTCAACGGTATTA 939

DB 6004 ATCAGCAGAAATCTATTAAGATCACTACCTACCCGAACCATGTATTCAGAGATGACAG 6063
QY 940 ACAGGTTTACATCTGCAGAGAGTCCGGTTAAACGTGCACCTTAATTCAGAGAAATT 999
DB 6064 CGAAGTACTCTGACCCAAAGTTTGTGAGCTGTGTGTAACTATCTGCAAGAGAAATT 6123
QY 1000 ACCCTCAGTACGAGCTTATTTATTAACAGATGATGATGATGATGATGATGATGATG 1059
DB 6124 ACCCGAGGTAGCATCTTTATCAATACAGACGACGATGATGATGATGATGATGATG 6183
QY 1060 ACCGCGCATGCTGCTGTAGATACAGCACTTTTGTGCGGCTAACTGAGAACTACC 1119
DB 6184 ACCGAGACAGTCCGCTGTGATATGCAACTTTTGTGCGGCTAACTGAGAACTACC 6243
QY 1120 CAAAGAGCTATGCTATTTGACGACGACGATGATGATGATGATGATGATGATGATG 1179
DB 6244 CGAAAGAGCAGAGTATGAGCCCAACATCCGAGTGGGTTCATCAGCATGACAG 6303
QY 1180 ATACATTAACAATGATTTGGCTGACGCTACTTAAAGAAATTCACAGTACCAATGC 1239
DB 6304 ACAGGTTGCAAAACGTGCTCATTTGCGGCTAAAGAACTGCAACGTACCAAAATGC 6363
QY 1240 GAGATTAACCTGCTTATGATTTGCGGCGCATTTATGTTGATTTCAAGAAATACGAT 1299
DB 6364 GTGACCTGCCAACACTGACCTGACGACATTCACAGTTGATGATGATGATGATGATG 6423
QY 1300 GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
DB 6424 GCAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6483
QY 1360 TTACGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419
DB 6484 TTACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6543
QY 1420 ATATATTAACCGGTTGACAGAGATACCAATGATCAATTCGATGATCTTAAAGAGAG 1479
DB 6544 ATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6603
QY 1480 ATGTCAAAGTATATCCCGGACGAAACATACAGAGAGAGCGGCTAAGTGCAGATATTC 1539
DB 6604 ACCTGAAAGTATACCTGCGACGAAACAGAAACAGAAACAGAAAGTATATC 6663
QY 1540 AGGCTGAGATCCCTGCTGACCGCTTACCTTTGCGGAGATCCATCGGAAATATGATC 1599
DB 6664 AAGCCGCAAGACCTTGGGACCGCTTACCTTATGCGGATCCACCGGATATGATGCGCA 6723
QY 1600 GACTGATGCGGCTGCTTGCCTCAATATCCATCTCTTGCATGATGACGCGAAAGATT 1659
DB 6724 GCTTACAGCGCTTTTGTATACCAACATTCACACGCTTTTGAATGATGCGGAGAGACT 6783
QY 1660 TTGATGCGATATGCTGTAACATTTTCCACGCGGACCCAGTATTTGAAACGGAATGC 1719
DB 6784 TTGATGCGATATGCGTAACATTTTCCACGCGGACCCAGTATTTGAAACGGAATGC 6843
QY 1720 CGTGTGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1779
DB 6844 CTTGTTGCAAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6903
QY 1780 ACTTATGATGACCAACCGCTTTTGAATTTGATGATGATGATGATGATGATGATGATG 1839
DB 6904 ACTTATGATGACCAACCGCTTTTGAATTTGATGATGATGATGATGATGATGATGATG 6963
QY 1840 CTGTGACCTACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1899
DB 6964 CCAACCATCTGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7023
QY 1900 TCTTAAAGCTGTTTGTGCAACACACTGATCAATATGATGATGATGATGATGATGATGATG 1959
DB 7024 TCTTAAAGCTGTTTGTGCAACACACTGATCAATATGATGATGATGATGATGATGATGATG 7083
QY 1960 AACGTTTAAACAGTACGAGTGGCGGCTTATTCGCGAGAGATTAATATGATGATGATG 2019

Db 7084 AGCGCTTAAAACTGCCAAATGTGACGACATTTATCGGCGACGACAACTTATACAGGAG 7143
Qy 2020 TCGTCTCCGACACTTGTATGGGAGAGATGGCCACTTGGTGAACATGAAATAAA 2079
Db 7144 TAGTATCTGACAAAGAAATGCTGAGAGGTGTCCACTTGGCTTCAACATGAGGTTAAGA 7203
Qy 2080 TTATTTGATGACATTATGGTATCAAGACCTTACTTCTGTGGGGATTTATCTGTGG 2139
Db 7204 TCATTGACGACATCGGCGAGAGACACTTACTTCTGGGTGGATTCATCTTGGACAG 7263
Qy 2140 ACCAGATTAACAGGACAGCTCTCAAGTGTGACAGCCCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7264 ATTCTGTTAATCTCAACAGGTGTCTGTGGCGGACCTCTTGAAGAGCTGTTTAAGTTGG 7323
Qy 2200 GAAACCAATTCGACGATGATACCAAGACTGAGACCGGCGCGGACCTGACATGATG 2259
Db 7324 GTAAACCGCTCCAGCCGACGATGAGCAAGACGAAAGACGAGCGCTGTGTAGATG 7383
Qy 2260 AAGCAATGCGATGAAACAGAAATTTGAAATTACGAGAGATTGAAAGCCGTAGATCA 2319
Db 7384 AAACAAAGCGGTGTTAGAGATGATTAACAGACACTTACAGTGGCGGTGGACATC 7443
Qy 2320 GATACGAGATCACTACGCGAGCCGTGATCATCTCTCTGTCAAGCTTAAAGCGG 2379
Db 7444 GGTATGAGGTAGACAACTACACCTGTCTGTGGCAATTGAAACTTTTGGCCAGACCA 7503
Qy 2380 TTAAGAACTTCAAGACATTAAGAGGAGCCCAATCACTCTTACGCGTGAAGCTTAATAG 2439
Db 7504 AAAGGCAATTTTAAAGCATCAAGAGGAAATTAAGCATCTTCAAGGTCTTAAATAGT 7563
Qy 2440 TGAAGTATGAA-----CAGCAGCTTACCCAGCGGAGATGTTTCA 2482
Db 7564 CAGCATAGTACTTTCATCTGACTAATACCAACACACACACACTGATAGAGATCT 7623
Qy 2483 TACCTCAGCTGAATTTTCACCAAGTTTACCTTAATCGATGCTTAAAGATCA 2542
Db 7624 TTAATATCTCTGGCGCGCCCTTCCAGCCCACTGCTGAGGCGCGGAGAA 7683
Qy 2543 AACCTCTAGGCGCGGTGAGGCGGCTTGGCCCGCTGGCTGCTCAATTCGAAGAT 2602
Db 7684 GGAGCAGGCGGCGCGGATGCTGCGCCGCAATGGCTGCTTCCAAATCCGCAACTGA 7743
Qy 2603 CTTAAGAGTCAATAGTCACTTGAATTTCAAAAC--GATCACTTACCGCCGAG 2660
Db 7744 CCAGCGCGTCACTGCTTCACTGATTTGACAGGCACTAGACTTAAACCCACCCAC 7803
Qy 2661 GTCCACCGCCAAAGAAAGAAAGTGTCTTAAAGCCAAACCTTACGCTTAAAGAA 2720
Db 7804 GCCCGCCCGCGCCAGAAAGAGGCGCCAAACCAACCGAAGCCGAAAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGAAAC---GCAAGCTTAAACAGGAAAGCAAGCTA 2777
Db 7864 AAACACAGAGAAAGAAAGAAAGCACTGCAAAACCCAAACCCGAAAGAAAGACGCTA 7923
Qy 2778 TGTGTATGAAGTTGGAGTGCAGCAAGCACTTCCGATC---ATGCTGAAGCCCAAGTGA 2834
Db 7924 TGGCACTTAAGTTGAGGCGGACAGACTGTTCGAGCTCAAAAATAGAGAGGAGATGTCA 7983
Qy 2835 ATGATATGCTGCTGTGTGAGGAGAGGCTGATGAACCACTTCAAGCTTGAAGAAAAA 2894
Db 7984 TCGGCGACGACATGCGCATGAAAGAAAGTAAATGAACCACTTCCACGTGAAGAACTA 8043
Qy 2895 TTGATTAAGAGCAATTAAGCGCGCTGAAATTGAAGAGCTAGCATGACATTTGGAGT 2954
Db 8044 TTGACACCGCTGTGTATCAAACTCAAAATTCACCAAGTCTGACATTAAGCAATGAGT 8103
Qy 2955 ACGGAGAGTTTCCAGAAATGAAATCAGACCGCTGACATGACAGGCAAGAACAC 3014
Db 8104 TCGCAAGTTCGCGGTCAACATGAGAGGCGCTTACCTTACACAGTGAACCCCTG 8163
Qy 3015 CGGCTTCTTCAATGCGCAACAGCGCGAGTCCAGTATGAGATGAGAGATTTACCGTAC 3074
Db 8164 AAGGCTTCTCAAACTGCGACCAAGGAGCGGTGAGTATGAGAGGAGATTTACATCC 8223

Qy 3075 CGAGAGAGTGGCGGAAAGGCGACAGCGGAAGACGATCTTGGACAAAGAGGACAG 3134
Db 8224 CCCCGAGTAGAGAGGACAGAGACAGTGTCTGCTCAATATGATTAATCAAGCCCGG 8283
Qy 3135 TTGTGGCTATTTCTTGAAGGTGCAATGAGGCGCGGTACAGCGCTTTCACTGTCTCA 3194
Db 8284 TTGTGGCATAGTCTTGGAGGCGGTGATGAGGGAACAAAGACCGCTTTCTGGTGTCA 8343
Qy 3195 CTTGAACCAAGAGGAGTACATTAAGGATACCCCGAAAGGTTCTGAACCGTGT--- 3251
Db 8344 CTGGAAATGCAAGAGGAGAAATCAAGACACCCCGAAGGGAACAGAAAGTGTGTG 8403
Qy 3252 -----CACTGATTAAGCGGCTATGCGGCTTTGAAATGTCAGCTTCCATGCGCAAC 3305
Db 8404 CTGACCACTGTGACAGGCTATGTCTTCTTGGAAAGCTAGCTTTCCATGCAATCCGC 8463
Qy 3306 CACCGTGTCTATTCATCTGACGCGAGAACGAACTTGAAGTGTCTGAAAGAAAGTCTG 3365
Db 8464 CGCCACATGTCTACCCCGGAAACCATCAGAGGTCTGACATCTGAAAGAAAGTCTGA 8523
Qy 3366 ACAATCCAAATTAAGACACGCTGTGAGAAAGTCTTGAAT---GTCCATCAAGCCGCG 3422
Db 8524 ACCAGAGCCCTAGACACCTGTCTCAACGCAATATTCGCGGTGATGTCTCGGCAAGA 8583
Qy 3423 CCAACGAAAGATTAACGATGACTTCACTGACGAGCCCTTACCTGGGCTTGGCCGCT 3482
Db 8584 GTAAAGAGGCTCATCTGACACTTTTACCTTGAACGCGCTTCTTGGGCAATGCTGT 8643
Qy 3483 ATTGACGACACTCAACGCGGTGTTTCAAGCCCAATTAATTAAGAACTGTGGAGCAAT 3542
Db 8644 ACTGTACACTATCTGAACCGGTCTTGAAGCCGATTAAGATGAGAGAGTCTGGAGTGAAG 8703
Qy 3543 CTGATGATGATCGATTAAGATTCAGAGTCTGGGCAATTTGGCTCAATAGGAGGCA 3602
Db 8704 CGGAGCAACACATAGCATAAGACTTCCGCGCAATTTGATACCAACAAAGCGGAG 8763
Qy 3603 CTGGAGATGACCAAAATTCGTTTCAATGCTTTTGAACAGACATGATCAAGAGAG 3662
Db 8764 CAGCAAGCTCAATTAAGTACCGCTTACATGCTGTGACAGAGATATCTGTAAGAG 8823
Qy 3663 ACAGATGAGAAATATAGCTATCAGACATCTGACCTGCGCTGTGTCGCGCAAG 3722
Db 8824 GCACATGATGATGATCAATGATCAGACCTAGAGACGCTTGAAGAGCTTATGATCAAG 8883
Qy 3723 GGTACTTCTGTGATGATGATGCTCTTCAAGTGAAGTGAACCTGATCAAGAGCG 3782
Db 8884 GATACTTCTCTGCGAAAGTGTCTTCAAGGAGACAGCTGAACGCTTGAAGAGAGTA 8943
Qy 3783 GAGCATGAGAAATTCATGACCGGTGAGAAAGATGAGAGAAAGTTCGTAGAG 3842
Db 8944 GCAACTGACAGCTGATGACAAATAGCGCCGCAAGATTAACAAATTCGTGGAGCGG 9003
Qy 3843 AGAGTACTTGTTCACCCGCTCATGGAAGCTGTAAATGCGACCTTTACATCACT 3902
Db 9004 AAAATATGACCTTACTCTGCTTCAAGATTAAGATTCCTTGCACATGTATGACGCTC 9063
Qy 3903 TGAAGAGAGCTTGTGCGGATCAATTAACATGCAAGGCGCAGGCGCACAGCGTAAAGT 3962
Db 9064 TGAAGAAACAAACGCGCGCTACATCACTATGACAGGCGCGGACCGCACTTATCAT 9123
Qy 3963 CCTATCTGAGAGAGCGTACAGGCGAAGTATTAACCACTTTTGGCAAGAACGCTCA 4022
Db 9124 CTTATCTGAGAGAAATCACTAGAGAAAGTTTACGGAACCAATTCGGAAGAAACATTA 9183
Qy 4023 CTTACGATTAAGTGTGGCACTACAGCAGGTATCTGAGACGCGAAGCAAGATGA 4082
Db 9184 CGTACAGTGAAGTGGCGGATTAACAGAACGGAACCGTTACGACCGCTACCGAAATCA 9243
Qy 4083 ACGCTGACATTAAGCAAAACAGTGCATTTGCTTCAAGAGGACCAACGAAATGGGTCT 4142
Db 9244 CGGCTGACCGCATCAAGAGAGTGTGCTTAAAGAGGACCAACGAAAGTGGGTCT 9303

QY	4143	CCAACTGCGCGGAGCTTTATTATGAGCAACAGACCACTGAGTGCAGAGTAAATTTGCAACTTC	4202
Db	9304	TCAACTGCGCGGAGCTGATCGACAGACCGCCGACCAAGCCCAAGGGAATTTGCAATTTC	9363
QY	4203	CATTCCGCTTGACACCGCAAGTGTGCGCGGTTCGGTATAGCTCACACCGCTTACAGTCCGA	4262
Db	9364	CTTTCAAGCTGATCCCGAGTACCTGCAATGCTCCTGTTCGCCACGCGCCGAACTGATAC	9423
QY	4263	AGTGGTTCAAAAGGCAATCACCTCTCCACCTGACTGAATGCGACCAACTTGTGCAACGA	4322
Db	9424	ACGGCTTTAAACACATCAAGCTCCCAATTAGACAGACCACTGACATTTGCTCCACCA	9483
QY	4323	GAAATTTGGGGCTGCGAGCAGAACCCACAGAGAAATGATTAACGGGTCTACATCCAGA	4382
Db	9484	GGAACTAGAGGGGCAAAACCCGGAAACCACTGAATGGATTCATGAAACCGGTTAGAA	9543
QY	4383	ATTTTTCGTGGGGCGAAGAGGGCTGAGTACTATGAGGTTGACCATGAAACAGTCCAG	4442
Db	9544	ACTTCACCGTGCAGCCGAGATGAGCTCTGGAAATCATATGGGCAATCAAGAACAGTAAGG	9603
QY	4443	TCTGAGGCCCAAGAGTCTGCACAGAGAACCTTCACGAGTGGCCACAGAAATAGTACAGC	9663
Db	9604	TCTATGCCCCAAGAGTCTGCACAGAGAACCTTCACGAGTGGCCACAGAAATAGTACAGC	9663
QY	4503	ACTATTTATCATCCGGCATCTCAGTCTACACTGTGCAATTTGCTGTGTGTGTGTGCTTCTCTA	4562
Db	9664	ATTACTATCATCCGGCATCTCGTGTGTAACCACTTAAGCCGTGCATCAGCTGTGTGCGA	9723
QY	4563	TCTCTGTATGAGCACTGCATCATCAGCAGCTTGCATTCGCGAAAGCAAGAAAGAGACTGCCTGA	4622
Db	9724	TGATGATTTGGGGTATCTGTTCGACATTAATGTGCTGTAAAGCGCCGCTGATGCTTGA	9783
QY	4623	CGCCATACGCGCTTGACCGAACCGCAACGATACCAAGCAATTAGCGGTTTGTGCTGA	4682
Db	9784	CGCCATATGCGCTGACCGCAATATCCGATTCCAACTTGCTGCTGCACTTTTGTGCTGTG	9843
QY	4683	TTGCGGCCAACCAAGCTGAAACATTTGGAGAACTTTGAACACTGTGTGTTTAACAAC	4742
Db	9844	TTTAGTCCGGCTAAATGCTGMAACATTTCAACGAGAACATGATGTTATTTGTGTGCAAGCC	9903
QY	4743	AACGTTTCTCTGGGACAGTTGTGCAATTCCTCTGGGACGCGCTTGTAATTCGTGTCGCT	4802
Db	9904	AGCGGTTCTTTGGGATCGAGCTGTGTATCCTCTGGCGCGCTGTGTGTGTTTAATGCGCT	9963
QY	4803	GCTTTTTCATGCTGCATGCTTTTATTTATTTGTTGAGGCGTCTGCTGCGGAGGTAGACG	4862
Db	9964	GTTGCTATGCTGCTGCTGCTTTTATTTAGTGTGCGGCGCTTACTGCGGAAGTAAACG	1002
QY	4863	CCTTCGAACATGCGACCACTGTGCGCAATGTTCGGGGATTCGCTATTAAGCGTTGTGCTG	4922
Db	10024	CCTACGAACATGCGACCACTGTGCTCAATTTGCGAACAGATATACCTTAATAGGCACTTGTG	1008
QY	4923	AACGCGCAGGTTAGGCGCGCATTTAACCGTGAAGATCAAGGTGCTCATTCGGAATTTAAC	4982
Db	10084	AAAGGCGAGGTTAGGCGCGCTCAATTTGGAGTTATCTGTATATCTCTCGAGGTTTTC	1014
QY	4983	CTTCAACTTAACAGAGATGACGTGACCTTGCAATTTCAACACAGTCAATTCCTTCAACAG	5042
Db	10144	CTTCAACCAACAGAGATGATTAACCTGCAATTTCAACACTGTGTGCTCCCTCCCTTAAAG	1020
QY	5043	TTAAATGCTGCGGCTCCTCGAGTGCAGAGCATCTCTAAAGCGGATTTACATGCGCG	5102
Db	10204	TCAATGCTGCGGCTCCTTGGAAATGTACGCCGCGCTCACGCGACATTAATCTGCAAG	1026
QY	5103	TTTTTTGGCGGTGTATCCCTTTCAATTTGGGGAGGCGCAATGCTTCTGTGACATGTAGA	5162
Db	10264	TCTTTTGGAGGGGTATACCCCTTCAATTTGGGGAGGAGCAATGTTTTTTCGACAGTAGA	1032
QY	5163	ACACAACTAGTGAAGGCGTACACTGAGTTGCTGCACGACTGCACTATATAGATCAACGAG	5222
Db	10324	ACAAGCAATATGATGAGGCGTATCGTCAATTTGTCAATGATTTGGCGCACTGACCAACGCG	1038
QY	5223	TGCACTAAAAGTTCAACAGACTGCTTGAAGTCGCGCTGCGATATGATATACGGCAACA	5282

Db	10384	AGCGGATTTAAAGTGTGCATATCTGCCCGCATGAAAGTASGACTGCCGATTAAGTGTACGGGAAACA	10444
Qy	5283	CCAACGGCGACACTGTGATACGTTTGTCAATAGCCGTCAACGCCAGGTTCTCACGGGACCTGA	5342
Db	10444	CTACCAAGTTTCTTAAGATGTGTACGTGAAGGAGTGTACACACAGGAACGTCTAAAGACTCTGA	10503
Qy	5343	AGGTCAATGACAGGCGCGATATCAAGCCGCTTTTTCACCCCTTTTGACCTTAAGGTGTGTATCA	5402
Db	10504	AAGTCATAGCTGGAACCAATTTCAGCATTTGTTTACACACTTTCGATTCACAAAGTGTGTATCA	10563
Qy	5403	GAAAGGGGCTTGTTTCAACTAGACATTCCTCCCTGATATGAGCTATGAAGCAACGAGAGCGT	5462
Db	10564	ATGCGGGCTGTGTGTACACTTAATGACTTTTCCGGAATACGAGCGATGAAGAACGAGAGCGT	10622
Qy	5463	TCGGCGATTTCAAGCATCTTCGCTTGTATGTCTACAGACATAGTAGCCGCACTGACATAC	5522
Db	10624	TTGGAGACATTTCAAGCACTCTCTTGTACTAGCAAAACCTCAATCGGCACAGACAGACATTAA	10683
Qy	5523	GGCTGTGAAGCCTTCTGTCAAGAACATCCACGTCCCTTACACCCACAGCATATCAAGGT	5582
Db	10684	GGCTACTCAAGCCTTCCGCGAAGAAAGTGTATCTCCGTAACACGACGGCGCATCTGAGAT	10744
Qy	5583	ATGGAATGTGGAAGAACATCACTGAGAGACCCCTCGAAGAAACAGACCAATTGGATGAT	5642
Db	10744	TCGGAATGTGGAAGAAACACTCAAGGCGCGCCACTGTGAGGAACCGCCCTTTTGGGTGTGA	10803
Qy	5643	AAATTGGAATGTGAGCCTCTGTGCGAGCGTCTTAATCTGTCTTACGGGCAATCCCTATCTGA	5702
Db	10804	AGATTGCAAGTCATTCGCGCTTCGAGCGGTGTGATGTCTATACGGGAACATTTCCATTCTTA	10863
Qy	5703	TTGACATCCCTGATGACAGCTTTTGTGTGATCATCAGAAATCCACAAATTTTAAGAGTTA	5762
Db	10864	TTGACATCCCGAAACGCTGCGCTTATACGAGCAATCAATGACACCACTGCTCTCAACGTGA	10923
Qy	5763	GCATGCAAGTAAGACGACTGACTTATTTCTGAGACCTTGTGTGTCTTAACATTTACAT	5822
Db	10924	AAATGTATGTCAAGTGTGATGACTTATTTACGGGACTTTCGAGGAGATGTGCTAACCTTCAGT	10983
Qy	5823	ACAAAGCTGACAGGAGGAGGACATTTGTCTCAGTTTCACTTCCACATCCACAGACAGCTGTTTGA	5882
Db	10984	ATGTATCCGACCGCGAAGAGACATATGCGCTGTACATTCGACTTCGACACAGCAACCTTCC	11043
Qy	5883	AGGAAGGACACACATGTGACTCTCCGTAGGACGATTAACACTTACTTTTACGACATGCA	5942
Db	11044	AAAGATGCAAGTTTCAATGTCTCTGGAAGAAAGAACGGGTGACATTAACCTTCAGACCGCGA	11103
Qy	5943	GCCCAACAAGCAATTTTATAGTTTTCCTTATAGCGGCAAGAAAGTCCACCTGCATATGCGAAT	6002
Db	11104	GCCCAACAGGGAATCTTATTTGTATCCGTGTGTGAAGAAAGCAACATGCATGACGAAGAT	11163
Qy	6003	GTAACCAACCGGCGGACACATATATTTGAGAAACCATATAAAGTCAGACCAAGATTTCCAG	6062
Db	11164	GCAAAACAACAGCTGTATCATATGTGTAGACACCCCGCAACAAAATATGACAAAGATTTCAAG	11223
Qy	6063	CGGCAATTTCCAAACATCTTGGAACTGCTGTCTGCATCTGTTTGGGGAGCATATCC	6122
Db	11224	CCGCAATCTCAAAAATTCATGAGATGTGGCTGTTTGCCTTTTGGCGGGCGCTCGTCCG	11283
Qy	6123	TCATTTGTTTGAAGACTTATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6182
Db	11284	TATTAATTAATGAGCTTATATATTTTGTCTTGACATGATGTGTGTGTGTGTGTGTGTGTGTGT	11343
Qy	6183	GACTGACGGGGAACATCTGA - CATATGCGGTAAACCTGATGTACTTCCGAGGAAGCGTGT	6241
Db	11344	GACCGCTAAGCCCCATATGACCCGACACGCAAAATCTGATGTACTTCCGAGGAACGTATGT	11403
Qy	6242	GCATTAATGCAACGCGCGCT-----TGA	6268
Db	11404	GCATTAATGCAATCAGCGGTGTATTTATGATCCCGCTTACCGCGGCAATATATGACACACC	11463
Qy	6269	AAAACTGATGATTTTCCAGGAAGACATGTGATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6318

Db 11464 AAACTCGAGTATTTCGAGGAGCGAGTCAATGCTGCGAGTGT 11513

RESULT 4
US-09-367-764-7
Sequence 7, Application US/09367764
Patent No. 6583121
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6583121th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09367,764
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-367-764-7

Query Match 37.0%; Score 2397.6; DB 4; Length 1163;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;

QY 580 ACTGACGCTATGAAACGGGAGCGTATATTTCTATCGGAAACAGGCCAGGTCACTTC 639
Db 5707 ACTGCTAACCGGGGAGTAGGTGATATTTTCACGAGACAGGCCCTGGGCACTTGC 5766
QY 640 AACGAAATCACTAGTCAATGTAACAAGAACTATATTTGATGGGCGCTCCATG 699
Db 5767 AAAAAGATCGCTTGTGAGAAACGAGCTTACGAAACCGACTTGGAGGCAATGTTCTGG 5826
QY 700 AGAAGTATTACGCCCGCGCTCGATCTCGAAAGAGAAATGTTACGAAAGAACTGC 759
Db 5827 AAAGATCTACGCCCGCGCTCGACAGTCCGAAAGAGAACGCTCAAACTCAGGTACC 5886
QY 760 AATTATGCGCTCTGTAAGAAATAGAGAGGTATCAATCAGAAAGTAGAAAAATATGA 819
Db 5887 AGATATGCCCCCAGGAGCAACAAAGAGGTACAGTCTCGAAAGTAGAAAAACAGA 5946
QY 820 AAGCAATTACAGCGAGGAGCTATTTTGTGATTTGGGCACTATCTATCATCAGAAGTGA 879
Db 5947 AAGCATTAACCACTGAGGAGCTGTTTCAGGAGGCTAGGCTGTATTAATCTGCGCAGC--- 6003

QY 880 ATCTGTGAGTGTATACAGATCAATTATCTGTACCAATCACTACGTCAACGTAATTA 939
Db 6004 ATCAGCCGAATATGCTATAGATACCTACCTAACCAACATGATTTCCAGAGCTGATCCAG 6063
QY 940 ACAGGTTTACATCTGCAAGAGTCCGGTTAAACGTCAACTTAATTATCCAGAGAAAT 999
Db 6064 CGAACTACTGACCCAAAGTTGCTGTAGCTGTGTTTAAACAATATCTGATGAGAAAT 6123
QY 1000 ACCCTACAGTACCGAGTATGATATGATACAGATGAAATAGATGCGATCTGACATGAGG 1059
Db 6124 ACCGACGTTAGCATCTTATACATCACCAGAGATAGATGCTTACTTGGATATGATG 6183
QY 1060 ACGGCGCATCTGTCTGTCTAGATACAGCCACTTTTGTCCGGCTAACTGAGAACCTACC 1119
Db 6184 ACGGAGAGTGGCTTGCTTACTGATGATGCACTTTTGTGCGCCGCAAGCTTGAATTTCC 6243
QY 1120 CAAGAGACATAGCTATTTTGCAGCCAGATATGATCAGCCGTCCATGCTTATACAGA 1179
Db 6244 CGAAAGACAGAGATATAGACCCCAACATCCGAGTCCGTTCCATCAGGATGACAGA 6303
QY 1180 ATACATTCAAAAATGATATGCTGACGCTACTAAAGAAATTCGAACTTACCCAAATGC 1239
Db 6304 ACAGCTGCAAAAGCGTCTCATGTCCGAGCTTAAAGAACTGCAACCTCACACAAATGC 6363
QY 1240 GAGAAATTAAGTCTGTATGATTCGGGCGCATTTAAATGTTGATTTTCAAGAAATAGCAT 1299
Db 6364 GTGAATCTCCCACTGATGATCTCAGGACATTCACAGTTGAATGCTTTGAAAATATGAT 6423
QY 1300 GCAATGATAGTACTGGGATACCTTTGCGATTAACCTTATTCGGTAACTACAGAAACG 1359
Db 6424 GCAATGACGAGTATTTGGAGAGATTTGCCGAAAGCCAAATTAGATCACTAGTATTCG 6483
QY 1360 TTACGCAATATGATGACAAAGTGAAGGCGGAAAGCAGCAGCATTTGTTGCAATATTC 1419
Db 6484 TTACCGCATAGTGGCCAGACTGTAAGGCGCTTAAAGGCGCGCACTGTTCCAAAGAGCG 6543
QY 1420 ATTAATCTAAACCGTTGACAGAGATACCAATGATCAATTTGTCATGATCTTAAAGAG 1479
Db 6544 ATTAATTTGCTCCATTTGTAAGAAAGTCTTATGATATGTTGTCATGATGATGAAABG 6603
QY 1480 ATGTCAAAATTAATCTCCGCGACGAAACATACAGAGAGCGGCTTAAAGTGAAGTTATTC 1539
Db 6604 ACGTGAAGATTAACACTGTGACAGAAACACAGAAAGAAAGCCGAAATTAACAGTATAC 6663
QY 1540 AGGCTGCGATATCCCTTGTCTACCGCTTACCTTTCCGGATCCATTCGGAATTAATGCTCTA 1599
Db 6664 AAGCGGCAAGACCCCTGCGACCGCTTACCTATGCGGATCCACCGGAGTTAGTGCACA 6723
QY 1600 GACTGAATGCGGTGCTTCTGCAAAATATCAATCTCTTGCACATGTCCAGCGGAAGATT 1659
Db 6724 GCGTTACAGCCGTTTGTGTAACCAATTCACAGCTCTTTGACATGTGCGCGAGAGACT 6783
QY 1660 TTGATGCGATTAATGCTGAACATTTTCAACAGCGGACCCAGTATTTGGAACGACATGC 1719
Db 6784 TTGATGCAATATATGACAGAACATTTCAAGAGGATGACCGGTACTGGAAGACGATATCG 6843
QY 1720 CGTGTGTTAAAGAGGAAAGACAGCGTATTCGCTATTTGCGCGTTGATATCTTTGAGG 1779
Db 6844 CCTGTTCGACAAAGCCMAACACACGCTATAGGGGTTAAACCGGCTGATGATCTTTGAGG 6903
QY 1780 ACTTAGGTTGAGACCAACCGCTTTAGATTTGATAGAGGGGCGTTCGGCAATATCAAT 1839
Db 6904 ACTGAGTGTGACCAACCACTTACTCACTTGAATGATGATGCGCTTTGGAATATATAT 6963
QY 1840 CTGTGACACTTACCTACAGAAAGAGTTTAAATTTGTGTCATGATGAATCCGATATGT 1899
Db 6964 CCACCATCTGCCCCAGCGGTATCCGTTTCAAAATTCGGGGCATGATGAATCCGGAATGT 7023
QY 1900 TCTTAAAGCTGTTTGTCAACACACTAGTCAATATATCATGATTTGATGACAGATACTAGTG 1959
Db 7024 TCTTCAGCTCTTGTGTCAACAGTTCGATATGATGATGATGATGATGATGATGATGATG 7083
QY 1960 AAGGTTAAACAGCTCAGCGTGCAGCGGCTTATTCGCGAGATTAACATAGTATGATGATG 2019

Db 7084 AGCGGCTTTAAACGTCCTCAAAATGTGCAGATTTATCGGAGACGACAACTTTACACGAG 7143
Qy 2020 TCGCTCCGACACCTTGATGGCGAGAGATCGGCACTTGCTGAACATGGAATMAAA 2079
Db 7144 TAGATCTGACAAAGAAATGGCTGAGAGTGGCCACTGGCTCAACATGAGGTATAGA 7203
Qy 2080 TTATGTATGACGATTATTGGTATCAAGACACCTACTTGTGGGGAGATTATCCGTGG 2139
Db 7204 TCATTTGACGACATCGGCGACAGACCACTTACTTGTGGGTGATTCATCTTGCAG 7253
Qy 2140 ACCAGATPACAGGACAGCCTGACAGAGTCGACACCTCTTAAAAAGGCTTTTAAAGCTG 2199
Db 7264 ATTGGTTACTCCACAGCGTGTGGCGTGGGAGCCCTTGAAGAGGTGTTTAAAGTTGG 7323
Qy 2200 GAAACACCTTCCGCTCGATATACCCAAAGTCGACCGCGCGGCACTGCATGATG 2259
Db 7324 GTAAACCGCTCCAGCGACATGACMAAGACGAAGACGAAGACGCGCTCTGCTAGATG 7383
Qy 2260 AAGCAATGCGATGGAACGAATTGGAATTACGAGATTAAGGCGGTAGAAATCCA 2319
Db 7384 AAACAAAGGCGTGTGTTAGATGATTAACAGACACTTACAGTGGCGGTGACATC 7443
Qy 2320 GATACGATCATCTGCGACGCGCTGATCATCGTCTGTGTCCACGTTAGCCGAAAGCG 2379
Db 7444 GGTATGAGTGAACAATCAACCTGTCTGTGGCATTGAGAACTTTGGCCAGAGCA 7503
Qy 2380 TTAAAGACTTCAAGACATTAAGAGGAGCCCAATCCTCTTACGGCTGACCTTAATGG 2439
Db 7504 AAACAGCATTTCAAGCCATCAGAGGGGAAATTAAGCATCTCTACGGTGTCTTAAATAGT 7553
Qy 2440 TGAGTATGATA-----CAGGACCTTACCACCGGAGATGTTTCCA 2482
Db 7564 CAGCATGTATCATTTCACTCTAATACCAACACCAACACCATGAAATGAGATTTCT 7623
Qy 2483 TACCCTAGCTGAACCTTTCACACAGTTTACCCTAACATTCGATGCTTACCGAATCCA 2542
Db 7624 TTAAATGTCTGGCGCGCGCCCTTCCAGCCCACTGCATGTGAGGCGCGGAGAA 7683
Qy 2543 AACCTCTAGGCGCGCTGAGGCGCTTTGGCGCGCGCTGGCTGCTCAATGGAAT 2602
Db 7684 GGAAGCAGGCGCGCCCGATGCTGCTCCCAATGGGCTTCCCAATTCAGCAATGTA 7743
Qy 2603 CTTAGAGGTGATAGTCACTTGACTTTCAAAAC--GATCACTTAATCGCGCGAG 2660
Db 7744 CCAAGCGGTAGTGCCTTACTTGTGACAGGCACTAGACTTAAACCCACGCCAC 7803
Qy 2661 GTCCACCGCAAGAGAGAGAGTGTCTCTTAAAGCCAAAACCTTACTGAGCTTAAAAA 2720
Db 7804 GCCCGCGCGCGCGCAAGAGAGAGGCGCCAAAGCAACCGAAGCGAAGAAACCA 7863
Qy 2721 AGAAGCAGCAACCAAGAGAGCAAC--GCAAGCTTAAACAGGAAACGACAGCTA 2777
Db 7864 AAACACAGGAAAGAGAGAGAGCACTGCAAAACCCAAACCGGAAAGAGACGCTTA 7923
Qy 2778 TGTGTATGAATGTGAGTGGAGCAAGACATTTCCGATC---ATGTGAAGGCCCAAGTA 2834
Db 7924 TGGCATTAATGTGAGGCGCAGACGATGTCGAGTAAAAATGAGGACGGAAGTGTCA 7983
Qy 2835 ATGATATGCTGTGCTGTGCGAGAGAGCTGATGAACCACTCCACGTTGAGAAAA 2894
Db 7984 TCGGGACGACGCTGGCCATGGAAGAAAGATATGAACCACTCCACGTTGAAGAACTA 8043
Qy 2895 TTGATATGACCAATTAGCGCGCGGTGAATTGAAGAGGCTAGCATGTGACATTTGGAGT 2954
Db 8044 TTGACACCCCTGTCTTCAAAAGTCAAAATTCACCAAGTCTGACATACACATGTGAGT 8103
Qy 2955 ACGGCGATTTCCCGACAGATGAATGACAGCTGCAATACACCGGACCAAAACAC 3014
Db 8104 TCGACAGTTGCGGTCAACATGAGAGTGAAGGCTTACCTTACACCAAGTAAACCTTG 8163
Qy 3015 CGGGCTTCTAACATGCGACCAAGCGCGAGTCCAGTATGAGATGAGATTTTACCGTAC 3074

Db 8164 AAGGTTCTAACACTGACCAACGAGCGGTGACGATATAGTGAGGACGATTTACATCC 8223
Qy 3075 CGAGAGAGTGGCGGGAAGGAGACGCGAAGACCGATCTGGAACAAGAGGACAG 3134
Db 8224 CCGCGAGTAGAGGAGAGAGAGAGTGTGTCTCGATTAATGATTAATCTCAGCGCG 8283
Qy 3135 TTGTGCTATGTTCTTAGAGGTGCAATGAGGACGCGTCAAGCGCTTCAAGTGTCA 3194
Db 8284 TTGTGCGATATGCTCGAGAGGCGGTGATGAGGAAACAAACCGCCCTTTCGTCGTA 8343
Qy 3195 CTGGAACCAAGAAAGGAGTCACTTAAGGATACCCCGAAGTTTGAACCGTGT--- 3251
Db 8344 CTTGAATACCAAGGAGAAACATCAAGCAACCCCGAAGGAGACGAAGAGTGTGTG 8403
Qy 3252 -----CACTAGTTACAGCGCTATGCGTGTTCGAAATGACGTTCCATGGAACAAC 3305
Db 8404 CTGCACACCTGTCACGCGCATGTGTGTTGAAACGTGAGCTTCCATGCAATGCC 8463
Qy 3306 CACCGTGTCTATTCACTGACCGCAAGAACAACTCGACGTCTTGAAGAGACGTCG 3365
Db 8464 CGCCCATGTCTACACCGCGGAACATCCAGAGCTCTGCATCTCGAAGAGACGTA 8523
Qy 3366 ACAATCCAAATTACGACAGCTGTGAGAACGTTTGAAT--GTCAATACGCGGC 3422
Db 8524 ACCACGAGGCTTACGACACCTGCTCAACGCCATATTGCGGTGCGGATCTCGGCA 8583
Qy 3423 CCAACGAACATTAACGATGACTTCACTGACCGCTCTCACTGAGGAGTCTGCGCGT 3482
Db 8584 GTAAAGAACGTCATCGAGACTTACCTTGAACAGCCGTACTTGGGACATGCTGT 8643
Qy 3483 ATTGCAACACTCAACGCGTGTTCAGCCCAATTAATAATTGAAACGTGTGGACGAAT 3542
Db 8644 ACTGTACCACTATCGAACCGTGTTAAGCCGATTAAGATCGAGAGCTCTGGGATGAG 8703
Qy 3543 CTGATGATGATGATTAATGATTCAGGTCTGCGCAATTCGCTTACATCAGCAGCA 3602
Db 8704 CGGACGACCAACCATTCGATACAGACTTCCGCGCGGATTGGAATAGACAAAGCCGAG 8763
Qy 3603 CTGCGATGTCACCAATTCGTTTACATGTTTTCGACGACGACATGACATCAAGAG 3662
Db 8764 CAGACACTCAATTAATTAACCGCTTACATGTGCTCGAGCAGATCATCTGTCAAGAG 8823
Qy 3663 ACAATGAGAAATAGCTATGACACATCTGGAACCTCGCTGTCTTGGCCACAAAG 3722
Db 8824 GCACATGATGATCATGAATGACGACCTCAGGACCGGTGAAGGCTTAAGTACAAAG 8883
Qy 3723 GGTACTTCTGTAGCTCAATGTCTTCAGGTGACATGTAACTGATCAAGCGC 3782
Db 8884 GATACTTCTCTCGCAATGTCTTCAGGGACAGCGTAAACGTTAGCATAGCGAGTA 8943
Qy 3783 GAGCATGTGAATTCATGCAACCGTGAGAAAAAAGATCAGAGAGATTGTGGTGAAG 3842
Db 8944 GCAACTGAGAACGTCTATGCAATAGGCGCGCAAGATTAACCAAAATGTGTGGAGCGG 9003
Qy 3843 AGGAGTACTGTTCCACCGCTCATGGAAGGTGTGAAGTGCACGTTTACGATCAT 3902
Db 9004 AAAAATATGACATCACTCCGCTTACGATGAAGAGATTCTTGCACGATGAACGCTC 9063
Qy 3903 TGAAGAGAGCTGTGCGGATCAATACATGACAGGCGCAACGCGCTATTAAGT 3962
Db 9064 TGAAGAAACAAACCGCGCTTACATCTATGACAGGCGCGGACCCCAAGCTTATCAT 9123
Qy 3963 CTTATCTGAGAGAGCTCAGGCGGAATGATCAATTAACCACTTCTGGGAGAAAGCTCA 4022
Db 9124 CTTATCTGAGAGATCATCAGGAAAGTTTACGGAAGCACTTCGGGAAAGAACATTA 9183
Qy 4023 CTTAGCAATGATGAGTGGGACTACAGCAGGTATCTGAGACCGAAGCAAGATGA 4082
Db 9184 CTTAGAGTCAAGTGGGAGTTTACAGACCGGAACGTTTACAGACCGGATACGAATCA 9243
Qy 4083 ACGGCTGCATTAAGCAAAACAGTCAATGCTTCAAGAGCGACCAAGCAATGGTCT 4142
Db 9244 CGGGCTGCACCGCATCAACAGTGTGCTGCTTATTAAGACGACCAAGCAAGTGGTCT 9303

QY 4143 TCACTCCGCGATCTTATTAAGGACACAGACCACTCAGTGCAGATTAATTCACATTC 4202
DB 9304 TCACTCCGCGGACTGATTCAGACACGCGACCAACGCCCAAGAAATTCATTC 9363
QY 4203 CATTCGCTGACACCGACAGTCTCCGCTTCCGTTAGTTCACACGCTTACAGTACGA 4262
DB 9364 CTTCACAGCTATCCGAGTACCTGATGCTCTGTTGCCACGCGCGAAAGTATAC 9423
QY 4263 AGTGTAAAGGATCAACCTCCACCTGACCTGCAATGCGAACCAATTCGTCGACAAAGA 4322
DB 9424 ACGGCTTAAACACATCAAGCTCCCAATTAGACACAGACCATCTGACATTCCTCACACCA 9483
QY 4323 GAAATTTGGGCTGCGACAGACAGACAGAGATTAAGATTAAGGATTCATCCAGA 4382
DB 9484 GGAAGCTAGGGGCAACCCGGAACCAACATGATGATCATCGGAAACAGTTAA 9543
QY 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTACGTAAGGGTAAACCATGAACCATGAG 4442
DB 9544 ACTTCAACGCTGACCGAGATGGCTGGAATACATATGGGGCAATCAAGAACAGTAAAGG 9603
QY 4443 TCTGGGCGGAGAGTCCGACACAGCGACCAATGATGCGCGATGATCATCATCC 4502
DB 9604 TCTATGCCAAGAGTCTGACACGAGACCTCAGAGATGGCAGACAAATAGTACAGC 9663
QY 4503 ACTATATCATCGGATCAGTCTACATGTCATTTGTCTGTGTGTGTGTGTGTGTGTGT 4562
DB 9664 ATTTACTATCATGCACTCTGTGTACCATCTTACCGGTGCGATCAGTGTGTGTGTGT 9723
QY 4563 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4622
DB 9724 TGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9783
QY 4623 CGGCAATAGCGCTTGCACGAGACGAGATCCAGACATTAAGCGTTTGTGTGTGTGTGT 4682
DB 9784 CGGCAATAGCGCTTGCACGAGACGAGATCCAGACATTAAGCGTTTGTGTGTGTGTGT 9843
QY 4683 TTCCGCAACCAACGCTGAAACATTTGGAAACCTTTGAACCATGTGTGTGTGTGTGT 4742
DB 9844 TTAGGTGCGCTATGCTGAACATTTACCGAGACATGATGTTATTTGTGTGTGTGTGT 9903
QY 4743 AACGCTTCTGTGGCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4802
DB 9904 AGCGTGTCTGTGGGTCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9963
QY 4803 GCTTTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4862
DB 9964 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10023
QY 4863 CTTTGAACATGCGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4922
DB 10024 CTTTGAACATGCGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10083
QY 4923 AACGCGAGTGAACGCGCACTTAACTGTGAGATCAGGTGTGTGTGTGTGTGTGTGT 4982
DB 10084 AAAGGCGAGGTGAACGCGCTCAATTTGAGATTAATGTGTGTGTGTGTGTGTGTGT 10143
QY 4983 CTTTGAACATGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5042
DB 10144 CTTTGAACATGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10203
QY 5043 TTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5102
DB 10204 TCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10263
QY 5103 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5162
DB 10264 TCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10323
QY 5163 ACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5222
DB 10324 ACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10383

QY 5223 TCGCAATTAAGTTCACAGACGCTCTGAAAGTCCGCTCGTATAGTATACGCAACA 5282
DB 10384 AGGGATTAAGTTCATCTCTCCGATGAAAGTATGATGATGATGATGATGATGATGATG 10443
QY 5283 CCACCGGCGACCTGATATGTTGTCAATGCGCTACGCGAGTTCTTACGCGACCTGA 5342
DB 10444 CTACAGATTTCTAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10503
QY 5343 AGGTATGACAGGCGGATATACGCGCTTTTACCTTTGACCATTAAGTGTGTGTGTGT 5402
DB 10504 AAGTATGATGTGACCAATTTTACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10563
QY 5403 GAAAGGCGCTTTTATCACTACATCTCTGTGATGAGTATGAAACAGAGGCT 5462
DB 10564 ATCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10623
QY 5463 TCGGATATTAAGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5522
DB 10624 TTGAGACATTCAGATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10683
QY 5523 GCGTGTGAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5582
DB 10684 GCGTGTGAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10743
QY 5583 ATGAAATGTGGAAGAACATCAGAGACCCCTGCAAGAAACAGACCATTTGTGTGT 5642
DB 10744 TCGAGATGTGGAAGAACATCAGAGACCCCTGCAAGAAACAGACCATTTGTGTGT 10803
QY 5643 AAATTTGAATGAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5702
DB 10804 AGATTTGAGATCAATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10863
QY 5703 TTGACATCCGAGACGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5762
DB 10864 TTGACATCCGAGACGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10923
QY 5763 GCTGACAGTACAGACGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5822
DB 10924 AATGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10983
QY 5823 ACAAGCTGACAGGAGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5882
DB 10984 ATGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11043
QY 5883 AGAAGGACCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5942
DB 11044 AAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11103
QY 5943 GCCCAGAGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6002
DB 11104 GCCCAGAGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11163
QY 6003 GTAACCAACGCGCGACACATTAATGTGAGAACCAATTAAGTGTGTGTGTGTGTGT 6062
DB 11164 GCAACCAACGCGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11223
QY 6063 CGGAGTTTCCAAACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6122
DB 11224 CGGAGTTTCCAAACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11283
QY 6123 TCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6182
DB 11284 TATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11343
QY 6183 GACTGACGCGGACACTGA-CATAGCGTAAACATGTGTGTGTGTGTGTGTGTGTGTGT 6241
DB 11344 GACCGTACGCGCGGACACTGA-CATAGCGTAAACATGTGTGTGTGTGTGTGTGTGTGT 11403
QY 6242 GCATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6288
DB 11404 GCATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11463
QY 6288 AAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6318

Db 7084 AGCGCTTAAAACTGCAAAATGTGCAAGCATTTATCGGCGACGACAACATTTATACAGGAG 7143
Qy 2020 TCGTCTCCGACACCTTTAGTGGGAGAGATGCGCACTTGGCTGAACATGAGATGAAAA 2079
Db 7144 TAGTATCTGACAAAGAAATGGCTGAGAGGTGTGCACTGCGCTCAACATGAGGGTTAAGA 7203
Qy 2080 TTATGATGCAAGTTATTTGGTATCAAGACCTTACTTGTGGGGATTTATCTGTGG 2139
Db 7204 TCATGAGGCAAGTCAATCGCGAGAGACCACTTACTTGTGGGTGATTTCACTTGGCAAG 7263
Qy 2140 ACCAGATTAACAGGCAACAGCTTCGACAGTTCGACAGCCCTCTTAAAAAGGCTTTTAAAGCTTG 2199
Db 7264 ATTCGGTAACTCCACAGCGGTGTCCGTGGCGGAGCCCTTGAAGAGGCTGTTAAGTTGG 7323
Qy 2200 GAAAACCATTTGCGAATGATGATACCAAGATGCGAGCCGCGCGGCGGCACTGCATGATG 2259
Db 7324 GTAAACCGCTCCACCGCAGATGAGCAAGACGAAAGACGAGCGCTGTGTAGATG 7383
Qy 2260 AAGCAATGCGATGGAACAGAAATTTGGAATTAACGAGATTAAGAGCCGTAGAAATCCA 2319
Db 7384 AAACAAAGCGGTGTTTAGATGATTAACAGACACTTGAAGATGGCCGTGGCAATC 7443
Qy 2320 GATACGAGATCACTACTGCGAGGCGTGAATCATCGTCTCTGTGCAAGTTAGCCGAAGCG 2379
Db 7444 GGTATGAGGTAGACAACATCAACCTGTCTGTGGCATTGAACTTTTGGCCAGAGCA 7503
Qy 2380 TTAAAGAACTTCAAGAGCATTAAGAGGAGCCCAATACCTCTAGCGCTGACCTTAAATGAG 2439
Db 7504 AAAGGCAATTTCAAGCCATCAAGAGGGAATTAAGCATCTTCAAGGTGTCTTAAATAGT 7563
Qy 2440 TGACGTAGTAGA-----CAGCACCTTACCCACCGGCGGAATGTTTCCA 2482
Db 7564 CAGCATATGATCTTTCATCTGACTAATACCAACAACACACATGATAGAGATTTCT 7623
Qy 2483 TACCTCACTGAACCTTTCCACCAAGTTTACCCTAACAAATCCGATGCTTTCGAGATCCA 2542
Db 7624 TTAAATGCTCGGCGCGCGCCCTTCCAGGCCCACTGTGAGAGCGCGGAGAGA 7683
Qy 2543 AACCTCTAGGCGCGCGCTGAGAGCGCTTTCGCGCCCGCTGGCTGCTCAATGCAAGAT 2602
Db 7684 GAGGCAAGCGCGCGCCGATGCTGCGCCGCAATGAGGCTGGCTTCCAAATCCGCAACTGA 7743
Qy 2603 CTTAGAGAGTGCATGATCACTTGAATTTCAAAACAAC--GATCACCTAATCCGCGCCAG 2660
Db 7744 CCACAGCGGTGATGCTTCACTGATGAGGAGGCACTGAGCTTCAAAACCCACGCCAC 7803
Qy 2661 GTCCACCGCCAAAGAAAGAAAGAGTGTCTTAAACCAAACTTACTGAGCTTAAAAAGA 2720
Db 7804 GCCCGCGCGCGCCAGAAAGAGGAGGCGCCAAACCAACGAGGCGGAAGAAACAA 7863
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGCAAC--GCAAGCTTAAACAGGGAACGACAACTTA 2777
Db 7864 AAACACAGAGAGAAAGAAAGAAAGCACTGCAAAACCCAAACCCGGAAGAGAGCTTA 7923
Qy 2778 TGTGTATGAGTTGAGTCCGACAGACATTTCCGATC--ATGCTGAACGCGCCAGTGA 2834
Db 7924 TGGCATTAAGTTGAGGCGCAGACAGCTGTTCGACGTCAAAAATAGAGACGAGATGCA 7983
Qy 2835 ATGATATGCTGCTGTTCGAGAGAGAGCTGATGAAACCACTCCACGTTGAGGAAAAA 2894
Db 7984 TCGGCGACGCACTGCGCATGGAAGAAAGGTAAATGAACCACTCCACGTGAAGAAACGA 8043
Qy 2895 TTGATATGAGCAATTAAGGCGCGCTGAATTTGAAGAGGCTAGCATGTGACGACTTGAAGT 2954
Db 8044 TTGACCACTGTGTCTATCAAAAGCTCAAAATTCACCAAGTGTCAACATACGACATGAGT 8103
Qy 2955 ACGGCGAGTTTCCCGAAGCATGAATGAGACAGCTGAGATACCGGCGACAAACGAC 3014
Db 8104 TCGCAGATGTCGGGTCAACATGAGAAATGAGGCGTTTACCTTACCGATGACACCTTG 8163
Qy 3015 CGGCTTTTACAACTGGGACCAACGCGGCAAGTCCAGTATGAGAAATGGAGATTTACCGTAC 3074

Db 8164 AAGGTTCTAACACTGGACCAACGAGCGGTGCTAGTATAGTGAAGGAGATTTACATCC 8223
Qy 3075 CGAGAGAGTGGGGGGGAAAGGCGACAGCGGAAACGCAATCCTGACAAACAGAGGACAG 3134
Db 8224 CCCGCGAGTGAAGAGGAGAGAGACAGTGTGTCTGCTCATTAATGATTAATCAAGCGCGG 8283
Qy 3135 TTGTGCTAATTTGTTCTAGAGAGTGCATAATGAGAGGACCGCTACGCGCTTTCAGTGTCA 3194
Db 8284 TTGTGCGCATATGCTCTCGAGAGGGGTGATGAGGAAACAAAGACCGCTTTTGTGCTGCA 8343
Qy 3195 CTGGAACCAAGAAAGGGGTGACATTAAGGATACCCCGAAAGTTTGAACCGTGT-- 3251
Db 8344 CCTGGAATGAGCAAAAGGGAAGCAATCAAGACAAACCCCGAAGGAGCAGAAAGTGTGTG 8403
Qy 3252 -----CATAGTTAACAGGCGTATGCGCTTGTGAATGTACAGTTCCATGCGCAAAAC 3305
Db 8404 CTGACCACTGTGTCAGGCGCATGTGCTTGTGGAAGGTAGAGCTTCCATGCAATCGCC 8463
Qy 3306 CACCGTGTGCTAATTCATCGACGCGAGAACGAACTGCAAGTGTGCGAAGAGACGTGCG 3365
Db 8464 CGCCCAATGCTTACACCCCGGAAACCATTCAGAGCTTCTGACATCTTCAAGAGAACGTGA 8523
Qy 3366 ACAATCCAAATTAACGACAGCTGTGAGAAAGTCTTGAAT--GTCCATCAACGCGCGC 3422
Db 8524 ACCAGAGCGCTTACGACACCTGCTCAACGCGCATATTCGCTGGGATGTCCGCGAGAA 8583
Qy 3423 CCAAAAGAAAGATTAACGATATTCACATGACACAGTCTTACTGTGGGATTTGCGCGCT 3482
Db 8584 GTAAAGAAAGGTACATGACATTTTACCTTGAACGAGCGGTACTTGGGACATGCTGT 8643
Qy 3483 ATTCAGACACTCAACGCGGTGTTTACGCGCCCAATTAATTTAGAAACGTGTGGAGCAAT 3542
Db 8644 ACTGTCAACATATGAAACCGGTCTTTAGCCGATTAAGATGAGACAGTCTGTGATGAG 8703
Qy 3543 CTGATGATGATGATTAAGATTAAGATTCAGAGTCTGCGCAAAATTCGCTCAATCAGCGAGCA 3602
Db 8704 CGGACGACAAACCATATGACATACAGACTTCCGCGCAATTTGATACGACCAAAAGCGAG 8763
Qy 3603 CTGCGGATGATCAACAAATTCGTTTACATGTTCTTTCGACCAAGACATATCAAGAGAG 3662
Db 8764 CAGCAAGCTCAATTAAGTACGCTGACATGTTGCTGACAGAGATCATCTGTCAAAAGAG 8823
Qy 3663 ACAGATGAGAAATGATCTTACAGACATCTGACACCTGCGCTGTCTTGGCCCAAG 3722
Db 8824 GCAACATGATGATCAATCAAGATCAAGACTCAAGACGCTGTGAGAGGCTTGTGATCAAG 8883
Qy 3723 GGTACTTCTGTAGCTCAATGTCTCTCAAGGTGACATGTTAACCGTCAATATCAAGAGCG 3782
Db 8884 GATATCTTCTCTCGGAAAGTGTCTCTCAAGGAGACGTTAGCATTAAGGAGTA 8943
Qy 3783 GAGCATCTGAGATTTATGACACCGTGAAGAAAGATCAGAGAGAGTTGTGCTAGAG 3842
Db 8944 GCAACTGACAACTCATGACAAATGCGCCGCAAGATTAACCAAAATTCGTGGAGACGG 9003
Qy 3843 AGAGTACTTGTTCACACCGGTCATGAGAAAGTGTAAAGTGTGACAGTTTACGATCACT 3902
Db 9004 AAAAATATGACCTTACCTCCGTTACGTTAAGAAATTTCTTGAACAGTATGACGCTC 9063
Qy 3903 TGAAGAGACGTCTGCGGGGTATCAATCCATGCAACAGCGCACGCGCCACCGCTATAGT 3962
Db 9064 TGAAGAAACAAACCGCGGCTATCAATCACTATGACAGCGCGGAGCCGATGCTATACAT 9123
Qy 3963 CTTATCTGAGAGAGCGTCAAGGCAAGTATCAATTAACCACTTCTGCAAGAACGTTCA 4022
Db 9124 CTTATCTGAGAGATTCATCAAGGAAAGTTTACGGAAGCCCAATCCGGAAGAAACATTA 9183
Qy 4023 CTTACGATGATGATGAGGAGCTTACAGACAGGATTCGTAAGACGCGGAACGAAAGTGA 4082
Db 9184 GTHACAGTGTGATGCGGATTTACAGAACGGAACGTTTACGACCTGTACCGAAATCA 9243
Qy 4083 ACGGCTGCACTTAAGCAAAACAGTGAATGCTTCAAGAGCGACCAACGAAATGGGTCT 4142
Db 9244 CGGCTGCAACGCGCATCAAGAGTGTGCTGTAAAGAGGACCAAAACGAAATGGGTCT 9303

QY 4143 TCACTGCGCGGATCTTATTAGGACACAGCACTGAGTGAAGTAATTTGACATTC 4202
DB 9304 TCAACTGCGCGGATCTCGATGAGACCGCAGCAACAGCCCAAGGAAATTTGATTC 9363
QY 4203 CATTCGCTGACACCGACAGTCTGCGGTTCCGTTAGCTCAGACGCTACAGTCA 4262
DB 9364 CTTTCAAGCTGATCCCGAGTACTGATGTTCCGTTAGCCCAAGCCGCAAGCTAGTAC 9423
QY 4263 AGTGGTTCAAGGATCACTCTCACTGACTGCAATGCCAACATTTGTCACAA 4322
DB 9424 ACGGCTTTAAACATCATGAGCTTCCAATTAGACACAGAACCATCTGATCTCAC 9483
QY 4323 GAAATTTGGGGCTCGAGACAGCAACAGCAATGATTAACAGGCTTACATCCAG 4382
DB 9484 GGAAGCTAAGGGCAACCCGAGCAACCACTGATGATCATGGAACAGGGTTAGAA 9543
QY 4383 ATTTTCTGTGGGGCGAAGAGGCTGAGTACGTATGGGTAAACCATGAACAGTCA 4442
DB 9544 ACTTCACTGACCGAGATGGCTGGAAATCATATGGGGCAATCAGAAACAGTTAAG 9603
QY 4443 TCTGGGCGCAGAGTCCGACACAGCCGACATGATGGCCGATGATCATCTCC 4502
DB 9604 TCTATGCCAAGAGTCTGACACAGAGAACCTTCAAGATGGCCACAGAAATAGTAC 9663
QY 4503 ACTATATCATCGGATCCAGTCTACCTGTCATTTGCTGTGTGCTGCTGCTGCTA 4562
DB 9664 ATTAATCATCTGCGCATCTGTGTACACATTTAGCCCTGCACTACGCTCTGTGCG 9723
QY 4563 TCTGTAGGACATCATCATCAGACCTTGCATCGCCAAAGCAAGAGACAGTCTG 4622
DB 9724 TGATGATTTGGGTAACGTGTTGACAGATTAATGCTGTAAAGCCGCGGTAGTCC 9783
QY 4623 CGCCATACGCGCTTGACACCGAACGAGTACCACAGATTAAGCGTTTGTGCTGA 4682
DB 9784 CGCATATGCGCTGCGCCCAATGCGGATTCACCTTCTGCTGCACTTTTGTGCTG 9843
QY 4683 TTGGGCAACCAAGCTGAAACATTTGAGAAACATTTGAAACATCTGTGTTTAAAC 4742
DB 9844 TTAGGTGGCTTAATGCTGAAACATTTACCGAGACATGATTTATTTAGTCAAGCC 9903
QY 4743 AACGTTTCTGTGGGACAGTTGTGATCTCTGTGAGCGCTTGTATTTCTGTTCC 4802
DB 9904 AGCGTTCTTCTGTGGTCAAGTGTGATTAACCTTGGCGCGTGTGTCTATAGGCT 9963
QY 4803 GCTTTTATGCTGATGCTTTTATTTATTTGTTGCAAGCGCTGCTGTGGAGGTA 4862
DB 9964 GTTGTCTATGCTGCTGCTTTTATTTAGTGTGTCGCGGCTTACCTGCGAAGGTA 10023
QY 4863 CTTTCAAGCATGGACAGCTGTGCAATGTTCCGGGATTCGGTATTAAGCGTTGCTG 4922
DB 10024 CTTAGCAACATGGACAGCTGTTCCAATGTGCAACAGATACGTTAAGCACTTGTG 10083
QY 4923 AACGCGAGGTTAGCGCGCACTTAACCTGAGATCAAGTGTGCTCATTCGGAATTA 4982
DB 10084 AAAGGCGAGGTTAGCGCCGCTCAATTTGAGATTATGTCATGCTCCGAGGTTTTC 10143
QY 4983 CTTTCAATTAACAGAGATGCTGACCTGCAATTTCCACACAGTATTTCTTACCA 5042
DB 10144 CTTTCAACCAACAGAGATGCTGACCTGCAATTTCCACACAGTATTTCTTCC 10203
QY 5043 TTTAAATGCTGGGGTCCCTGAGTGAAGGATCTCAAAAGGGGATTAACATGCGCG 5102
DB 10204 TCAATGCTGGGCTCTCTTGGAAATGTCAAGCCGCGCTCAGCGACATTAATCT 10263
QY 5103 TTTTGGCGGTGTACCTTTTCAATGTGGGAGCGCAATGCTTCTGTGACAGTGA 5162
DB 10264 TCTTTGAGGGGTATACCCCTTCAATGTGGGAGGAGCAATGTTTTTGGACAGTGA 10323
QY 5163 ACACACAATGATGAGGCGTATGCTGAGTTGCTCAAGTGCATTAATAGTACGAG 5222
DB 10324 ACAGCCAGATGATGAGGCGTATGCAATTTGATGATTTGCGCAGTACACGCGC 10383

QY 5223 TCGCAATAAAGTTACACAGAGTCTGTGAAGTCCGCTGCGTATGATATACGCA 5282
DB 10384 AGCGATTAAGTGTCAATCTGCGCCGATGAAGTAGAGTGTGTATAGTACCGGA 10443
QY 5283 CCACCGGCACTGATATCGTTGTCAATGCGCTCAGCCAGGTTCTCACGGAGCTGA 5342
DB 10444 CTACAGATTTCTAGATGTGATGTAACGAGTACACAGGAAACGTTAAAGACTGA 10503
QY 5343 AGGTCATAGAGGGCGATATCAAGCCCTTTTACCCCTTTGACCAATAGTGTCTA 5402
DB 10504 AAGTCATAGTGAACCAATTTGAGCATTTGTTACACCATTCGATCAAGAGTGT 10563
QY 5403 GAAAGGGCTGTGTTTACATTAAGCACTTCCCTAGATGAGGATGAACCAAGAGCT 5462
DB 10564 ATCGCGCTGTGTATCACTATGATCTTTCGGAATAGGAGATGAACAGAGCT 10623
QY 5463 TCGCGATATTCAAGCATCTGCTGTGATCTACAGATATGATGCCGACATGATAC 5522
DB 10624 TTGGAGCATTTCAAGCTTACCTCTTGTACTAGCAAAAGCTCATGCGCAGACAG 10683
QY 5523 GCGTCTGAAGCTCTGTGTCAAGATCAAGTCCCTTACCCCAAGCAGTACAGGCT 5582
DB 10684 GCGTACTCAAGCTTTCGCAAGAACTGTGATCTCCGTACAGAGGCGCATTTGAT 10743
QY 5583 ATGAATGTGAAGAAACAATCTCAGAGCACTCCCTGCAAGAAAGCAGCATTTGAT 5642
DB 10744 TCGAGATGTGAAGAAACAATCTCAGAGCGCCCACTCAGAGAAACGCCCCCTT 10803
QY 5643 AAATGAGTGAAGCTCTGCGAGCGCTTACTGTGCTTACGGGACATCTTATCTGA 5702
DB 10804 AGATTGAGTCAATTCGCTTGTGAGCGGTGAGCTGCTATACGGGAACATTTCTA 10863
QY 5703 TTGACATCCCTGATGAGCTTTTGTGATCATCAGAAATCAACAATTTTGAAGTTA 5762
DB 10864 TTGACATCCCTGAAAGCTGCTTTTATCAAGACATCAAGTGAACCATGCTTCA 10923
QY 5763 GCTGACAGTGAAGACTGTGATTTATCTCAGACTTGTGTGTTCTTAACATTAAGT 5822
DB 10924 AATGTGATGTCAAGTGTGATCTTATCAGCGGACTTCCGAGGATGCTTACCTGA 10983
QY 5823 ACAAGCTGACAGGAGGAGCAATGTTCAGTTCATCTCCATCTCAGCAGACTGTTTGA 5882
DB 10984 ATGATCTCAGCCGAGGAGCAATGCGCTGTATCTTGTGATGAGACAGCAACCTTC 11043
QY 5883 AGGAAGGACACACATGTATCGCTGAGGAGCAATTAACATTAATTTAGCATTCGA 5942
DB 11044 AAGAGTGAAGTTCATGTCTGAGAAAGAGCGGTGACAGTACTTGAAGACGCG 11103
QY 5943 GCCCACAAGAAATTTATAGTTTGTGCTATGCGGCAAGAGTCACTCGATGTGAAT 6002
DB 11104 GCCCAGGCGAACTTCAATTTGTATCGCTGTGTGTGAAGAGACATGTAGAGAT 11163
QY 6003 GTAAACCAAGGCGGACCATTAATGTGAGAACCAATTAAGTGAACCAAGAAATTC 6062
DB 11164 GAAACCAAGGCGGATCATATGTGTGAGCAACCCGCAAAATGTACCAAGAAATTC 11223
QY 6063 CGGCAAGTTTCCAAACATTTTGAACCTGTGTGCTGTGACCTGTTTGGGGGAGAT 6122
DB 11224 CGGCAATCTAAAACTTCAATGAGTGTGTGCTGTTTGTGCGGGGCGCTGTGCG 11283
QY 6123 TCAATGTTGAGGACTTATAGTGTGTGCTGTGAGCTTATGCTTATTAACACAGTAT 6182
DB 11284 TATTAATTAATGAGACTTATATATTTTGTGCTGTGAGCATGATGCTGACAGAG 11343
QY 6183 GACTGACGCGGACACTGA-CATAGCGTAAACCTGATGTACTTCCGAGAAAGCTGT 6241
DB 11344 GACCGTACGCCCAATGAGCCGACAGCAAACTGATGTACTTCCGAGAACTGATGT 11403
QY 6242 GCATTAATGCAAGCGCGCT-----TGACACT 6268
DB 11404 GCATTAATGCAATGCGGCTGTGATTAATGATCCCGTTACCGGGGCAATTAAGCAAC 11463
QY 6269 AAAACTGATGTATTTCCGAGGAAGCACAGTGCATTAATGTGTGAGTGT 6318

Db 11464 AAAAAGTCCAGCTATTTCCAGAGAAAGCGAGTGCATTAATGCTCGCAGATGT 11513

RESULT 6

US-08-801-263A-1

Sequence 1, Application US/08801263A

Patent No. 5811407

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Stimpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESS: Bell Seltzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 5811407th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,263A

FILING DATE: 19-FEB-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 11663 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 60..7559

FEATURE:

NAME/KEY: CDS

LOCATION: 7608..11342

US-08-801-263A-1

Query Match

Best Local Similarity 64.6%; Pred. No. 0;

Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

Score 2396; DB 1; Length 11663;

Db 5707 ACTGCTTAACCGGGTAGGTGGTACATTTTTCAGACGACAGCCCTGGGCACTTGC 5766

Db 640 AACAGAAATCAGTACGTCAATGTAACACTCAAGAACTTATTTGATGGGCGCTCATG 699

Db 5767 AAAAGATGCTCTTTCGCAAGACCACTTACAGAACCACTTGGAGCCCAATGTTTCG 5826

Db 700 AGAAGTATTACGCGCCCGCTCGATCTCGAAGAGAGAAATGTTTACAGAAAGAACTGC 759

Db 5827 AAAAGATCTACGCGCGGTGCTGACAGCTGAAAGAGAAAGCAAGCTCAAACTCAGTACC 5886

Db 760 AATTATGCGCTCTGAAGAAATAGAGAGATTCATCAGAAAGTAAATATGTA 819

Db 5897 AGATGATGCCACGAGAAAGCAAAAGCAGGTACCACTCTGAAAAGTAAAGCAACA 5946

QY 820 AAGCAATTACAGCGAGGAGACTATTTCTGATTTGGGACATATCTATCATCAGAGTGA 879

Db 5947 AAGCATTAACCACTGAGGAGCTTTTCAAGGGCTTACGATGTAATCTCTGCAAG--- 6003

QY 880 ATCTGTGAGTGTATACAGAGTCAATTTCTGTACCAATCTACTCTGCAAGGTATTA 939

Db 6004 ATACGCGAATGTCTATTAAGATCACTTACCGAACCATGTAATCCAGCACTGACAG 6063

QY 940 ACAGTTTACATCTGACAGGTCGGGTTAAAGTGCACCTTAATTTCAAGAAAT 999

Db 6064 CCAACTTCTGACCCAAAGTTTGTGTAGCTGTTTGAACAATCTGATGAGAAAT 6123

QY 1000 ACCCTACAGTACCGAGTATTTGATATACAGATGATGATGATGATGATGATGATG 1059

Db 6124 ACCGAGGTGATGATTTATATGATGATGATGATGATGATGATGATGATGATGATG 6183

QY 1060 ACGGCGATCTGCTGTCTGATATACAGCACTTTTGTCCGCTTAACTGAGAGCTACC 1119

Db 6184 ACGGACAGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 6243

QY 1120 CAAGAAGCATAGCTATTTGACAGAGATTAAGTACAGCGCTCCATCGCTATACGA 1179

Db 6244 CAAAAGACAGAGATTAAGAGCCCAACATCCGAGTGCCTTCCATCAGAGATGAGA 6303

QY 1180 ATACATTACAAATGATTTGCTGAGCTACTTAAAGAAATTCACACCTTACCAATTC 1239

Db 6304 ACAGCTTCAAAAGCTGTCTATGCTGCGGAGCTTAAAGAAATTCACACCTTACCAATTC 6363

QY 1240 GAGAAATTAAGCTGTCTTGAATTTGCGGCAATTAATGTTGATTTTCAAGAAATTC 1299

Db 6364 GTGAATCTCCAACTGAGTACGAGCACTTCAACGTTGAATGCTTGAAGAAATTCAT 6423

QY 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359

Db 6424 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6483

QY 1360 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419

Db 6484 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6543

QY 1420 ATATCTTAAACCGTTGCAAGAGATACCAATGATCAATTTGCTGATGATCTTAAAGAG 1479

Db 6544 ATATCTTAAACCGTTGCAAGAGATACCAATGATCAATTTGCTGATGATCTTAAAGAG 6603

QY 1480 ATGCAAGTATCTCCGCGACGAAACATACAGAGAGAGAGAGAGAGAGAGAGAGAG 1539

Db 6604 ATGCAAGTATCTCCGCGACGAAACATACAGAGAGAGAGAGAGAGAGAGAGAGAG 6663

QY 1540 AGGCTGAGATCCCTTCTACCGCTTACCTTTGCGGAGTCCATCGGAAATTAATGCTG 1599

Db 6664 AGGCTGAGATCCCTTCTACCGCTTACCTTTGCGGAGTCCATCGGAAATTAATGCTG 6723

QY 1600 GACTGATGCGGTGCTTTCGCAATATCAATCTCTTTCGACATCTCAGCGGAAGAT 1659

Db 6724 GACTGATGCGGTGCTTTCGCAATATCAATCTCTTTCGACATCTCAGCGGAAGAT 6783

QY 1660 TTGATGCAATTTATGCTTAAATTTCCACGCGGAGACCGGATTTGGAAGCGATTCG 1719

Db 6784 TTGATGCAATTTATGCTTAAATTTCCACGCGGAGACCGGATTTGGAAGCGATTCG 6843

QY 1720 CGTGTGTTGATTAAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1779

Db 6844 CGTGTGTTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6903

QY 1780 ACTTATGCTGACCAACCGCTTTAGATTTGATAGAGAGAGAGAGAGAGAGAGAGAGAG 1839

Db 6904 ACTTATGCTGACCAACCGCTTTAGATTTGATAGAGAGAGAGAGAGAGAGAGAGAGAG 6963

QY 1840 CTGTGACATCTACATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1899

Db 6964 CTGTGACATCTACATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7023

QY 1900 TCTTAACGCTGTTTGTCAACACATGATCAATATCATGATTTGCTAGAGAGATCTAG 1959

Db 7024 TCCACGCTCTTGTCAACAGATTCGTAATGCTGTTATGCGCAGCAGATATTTGAGG 7083
Qy 1960 AACGTTAACCACTGACCGCGCGCTCTATCGCGGAGAGTAACATATGATGATG 2019
Db 7084 AGCGGCTTTAAACCTCCAAAGTGCAGATTTATCGGAGAGCAACATTTATACAGAG 7143
Qy 2020 TCGCTCCGACACCTTGATGCGGAGAGATGCGCACTTGCTGAACATGAGATTAATA 2079
Db 7144 TAGATATCGAAGAAATGCTGAGAGTGTCCACTGCTCAACATGAGATTTAGA 7203
Qy 2080 TTATATGACGATTAATTGATCAAAAGCCCTAATTCTGTGGGGAATTAATCTGTG 2139
Db 7204 TCATATGCGAGCATCGGCGAGAGACCACTTACTTCTGCGGTGATTCATCTTGAAG 7263
Qy 2140 ACCAGATTAACGAGCAGCGCTGCAAGTGCAGACCTCTTAATAAGGCTTTTAAAGCT 2199
Db 7264 ATTGCGTTACCTCCACAGCGGTGCGGTGCGGACCCCTGMAAAGCTGTTTAAAGTTGG 7323
Qy 2200 GAAACCATTTGCGAGTATATCCAGACTGCGACCGCGCGGCGCATGATGATG 2259
Db 7324 GTAAACCGCTCCAGCGGACGATGAGCAAGACGAAGACGAAGACGCGCTTGTCTAGTG 7383
Qy 2260 AAGCAATGCGATGAGAACGATTTGAATTAAGGACGATTAAGTGAAGCCCTAGAAATCA 2319
Db 7384 AAACAAAGGCGGTGTTAAGATGATTAACAGACACTTGAAGTGGCGGTGCACTC 7443
Qy 2320 GATTCGAGATCATCTGCGAGGCTGATCATCACTCTCTGTCCAAGTGAAGCGAAGCG 2379
Db 7444 GGTATGAGTATGACACATCAACCTGTCTGTGGAATGAGAACTTTTCCACAGCA 7503
Qy 2380 TTAAAGACTTCAAGAGCATTAAGAGAGGCCCAATCACCTCTAGCGGCTGAATTAAG 2439
Db 7504 AAAGAGCATTTCAAGCATCAAGAGGGAATAAGCATCTTACGGTGTCTTAAATAGT 7563
Qy 2440 TGAAGTATGA-----CAGCAGCTTACCCAGCGGAGAAATTTTCA 2482
Db 7564 CAGCATGATCATTTCACTGATTAATACCAACACCAACCATGAATGAGATTTCT 7623
Qy 2483 TACCTTCACTGAATCTTCCACCAAGTTTACCCTAACATCCGATGCTTACCGAGATCCA 2542
Db 7624 TTAACATGCTCGGCGCGCGCTTCCAGCGCCCACTGCATGTGAGGCGCGGAGAA 7683
Qy 2543 AACCTCTTAGGCGCGCTGAGGCGCTTTCGCGCGCGCTGCTCAATGAGAT 2602
Db 7684 GAGGCGAGGCGCGCGCTGCTGCTGCGCAATGAGGCTGCTTCCCAATCCAGACCTGA 7743
Qy 2603 CTTAGAGAGTGCATGATCACTTGAACCTTCAACAC--GATCACTTAATCCGCGCAG 2660
Db 7744 CCAAGCGCTGAGTGCCTGATGATGGAAGGCACTGAACCTCAACCCAGCCGAC 7803
Qy 2661 GTCCACCGCCAAAGAAAGAAAGTGTCTTAAGCCAAAACCTACTCAAGCTTAAAGA 2720
Db 7804 GCCCGCGCGCGCGCAAGAAAGAGGCGCCAAAGCAACCCAAACCGGAAGAAACAA 7863
Qy 2721 AGAAGCAGCAGCAGCAGGAGCAAAAC--GCAAGCCTTAAACGAGGAAACGAACAAGTA 2777
Db 7864 AAACACGAGAGAAAGAAAGAGAGCACTTGCAAAACCCAAACCGGAAGAGAGCGGTA 7923
Qy 2778 TGTGTATGAAGTTGAGTGCAGCAAGACATTTCCGATC--ATGCTGAACGCGCAAGTGA 2834
Db 7924 TGGCACAATTAAGTTGAGGCGGACAGACTGTTGCACTCAAAATAGAGAGCGGATGCA 7983
Qy 2835 ATGATATATGCTGCTGTGTGCGAGGAAAGCTGATGAACCACTCAAGTGAAGGAAATA 2894
Db 7984 TCGGCGACGCACTGCGCATGAGAAAGTAATGAACCACTCAAGTGAAGGAACTA 8043
Qy 2895 TTGATTAATGAGCATTAAGCGCGCTGAATTAAGAAAGGCTAGCATGATGAGACTTGAAGT 2954
Db 8044 TTGACCACTCTGTGCTTCAAAAGCTCAAAATTCACCAAGTGTCTACATACAGCATGAGAT 8103
Qy 2955 ACCGCGAGCTTCCCAAGCAATGAAATCAGACGCTGAGTACACGAGCAACCAAC 3014

Db 8104 TCGCACAGTTGCGGCTCAACATGAGAGTGAAGGCGTTTCACTTACCACTGAACACCTTG 8163
Qy 3015 CGGCTTTCTACCACTGGCACACAGGCGAGTCCAGATATGAGAAATGGAGATTTAACGTAC 3074
Db 8164 AAGGTTCTACCACTGGCACACAGGCGGCTGAGTATGAGAGGAGATTTACATCC 8223
Qy 3075 CGAGAGAGTGGCGGGAAGGCGACAGCGGAAGACCGATCTTGAACACAGAGCGAG 3134
Db 8224 CCGCGGAGTAGAGAGGAGAGAGAGAGTGTGTCCGATTAATGATTAATCAAGCGCGG 8283
Qy 3135 TTGTGCTATTTGTTCTAGAGGTTGCAATGAGGCGACGCTTACGCGCTTCACTGATGCA 3194
Db 8284 TTGTGCGATATGCTCTCGAGGCGGTGATGAGGAAACMAAACCGCTTTCGCTGTCA 8343
Qy 3195 CTGGAACCAAGAAAGGGGTGACATTAAGGATACCCCGAAGGTTGAAACCGTGT--- 3251
Db 8344 CTTGAATAGCAAGGGAACAAATCAAGCAACCCGGAAGGAGACGAAGAGTGTCTG 8403
Qy 3252 -----CACTAGTTACAGCGCTATGCGTCTTTCGATGTCAAGTTCCATGCGACAAAC 3305
Db 8404 CTGACACACTGTGTCAGCGGCATGTGCTTGTGAAACGTGAGCTTCCATGCAATGCC 8463
Qy 3306 CACCGTGTCTATTCACTGACCCAGAACGAACCTGACGTGCTGAAAGAGACGTG 3365
Db 8464 CGCCCATGCTTCAACCCCGGAACCATCCAGAGCTCTGCACTCTGAAAGAAACCTGA 8523
Qy 3366 ACAATCAATTTGACACGCTGTGAGAGACGCTTTGAAT---GTCCATACGCGCGC 3422
Db 8524 ACACAGGCTTACAGACCTCTCTCAACGCAATGTGCGGTGCGATCTCCGAGAA 8583
Qy 3423 CCAACGAGACATTAACGATGATGATCACTGACAGTCCCTTACCTGGGTTTGCCTG 3482
Db 8584 GTAAAGAGACGTGTCAGAGACCTTAACTTGAACAGCCGCTTGGGCACTGCTGT 8643
Qy 3483 ATTGACAGACCTAAGCGCGCTGTTTCAAGCCCAATTAATAATTGAACGTGCGACAT 3542
Db 8644 ACTGTCACTACTGAACCTGTCTTGAAGCCGATTAAGATGAGAGTGTGGAATGAAG 8703
Qy 3543 CTGATATGATGATTAAGATCAGGTGCGGACCAATGCGCTCAATCAGGAGCA 3602
Db 8704 CGAGCGAACACATTCGATACAGACTTCCGCTGAGTTGATACACCAAGCGGAG 8763
Qy 3603 CTGCGATGTCACCAATTCCTGTTAATGATCTTTGACCAAGCACTGACATCAAGAG 3662
Db 8764 CAGGAGCTCAATTAATGATCCGCTTACATGTGCTGAGAGAGTCAATCTGTAAGAG 8823
Qy 3663 ACAGTATGAGAAATATGCTATGACGACATCTGGAACCTGCTGTGCGCAAG 3722
Db 8824 GCAACATGATGATCAATCAGATGAGCACTCAGAGACGCTGTAGAGGCTTGAACAAAG 8883
Qy 3723 GGTACTTCCGTAGTCAATGTCCTCCAGGTTGACAGTGAACGCTGATGACAGGCG 3782
Db 8884 GATACCTTCTCTCGGAGAGTGTCTCCAGGAGACAGCTGAAGGTTAGATAGCGAGTA 8943
Qy 3783 GACATCTGAGATTAATGACCGGTGAGAAAGATCAAGAGAGATTTGTGCTAGAG 3842
Db 8944 GCAACTCAGAACGTGATGACATGAGCCCGCAAGATTAACCAAAATGCTGAGAGCGG 9003
Qy 3843 AGAGATCTTGTTCACACCGCTCATGAGAAAGCTGTGAATGCTCAAGTTCATGATCACT 3902
Db 9004 AAAAATATGACTACCTCCGCTTACGCTGAAGAAATTCCTTGCACAGTGTAGACCGTTC 9063
Qy 3903 TGAAGAGAGTGTGCGCGGTATATTAACATGACAGGCGCAGCGCATGCTATTAAGT 3962
Db 9064 TGAAGAAACACCGCGGTATATCACTATGACAGGCGGAGACGCAATGCTATTAAT 9123
Qy 3963 CTTATCTGAGAGAGCTCAGGCGAAGTGAATTAACCACTTCTGCAAGAGAGTCA 4022
Db 9124 CTTATCTGAGAGATCATCAGGAAAGTTTACCGAAGCAACATCCGGGAAGAACTTA 9183
Qy 4023 CTTACGATTAATGTGTGCGACTACAGCAGATATGCTGAGACGCGAAGAGATGA 4082
Db 9184 CTTACGATGCAAGTGTGCGGATTAACAGACCGGAACCGTTTACGACCCGTAACCAATCA 9243

4083 ACGGCTGCACTAAAGCAAAACAGTCATTGCTTACAGAGCGAACCAACGAATGGGCT 4142
9244 CGGGCTGACCGGCATCAAGAGTGCGTCTATTAAAGGCAACCAAGAGTGGGCT 9303
4143 TCACATCCGCGGATCTTATTAGGCAACAGACCACTAGTGCAGTAATTGCACTTC 4202
9304 TCACATCCGCGGATCTGATCAAGACAGCGGACCAACGGGAAATGGCACTTGGC 9363
4203 CATTCGCTTGACACGCAAGTCGCGGCTCCGTTAGCTACACGGCTACAGTACAGA 4262
9364 CTTTAAGCTGATCCGAGTACTCTAGTGTCCGTTGGCCACGCGCAAGTAATAC 9423
4263 AGTGTCAAAAGGATCAACCTCTCACTGACTGACATGACCAATTTGTCAGAACGA 4322
9424 ACGGCTTTAAACATCAAGCTCCCAATTAGACAGACCAATGATGACTTGTCAACCA 9483
4323 GAAATTTGGGCTGCGAGACAGCAAGCAAGCAAGTAAGATTACAGGGTCTACATCAGA 4382
9484 GGAGACTAGGGGCAAAACCGGAAACCACTGAATGATCAATCGMAAACGCTTAGAA 9543
4383 ATTTTCTGTTGGGGGAGAGAGGCTGAGTACGTATGGGTACATTAACATCAAG 4442
9544 ACTTACCGTGCACCGAATGCGCTGGAATACATATGGGCAATACAGAACGATAGGG 9603
4443 TCTGGGCGCAGGAGTCCGACACGAGCGACCCACATGATGCGCGATGATCATATCC 4502
9604 TCTATGCCAGAGTCTGACACGAGAGACCTCAAGATGCGCAACGAATATGACACG 9663
4503 ACTATTAATCATGCGATCATCTACCTGATCTGATGCTGATGCTGCTCTGCTGTA 4562
9664 ATTAATCATATGCGCATCTGATGATACCATCTTACCGCTGATGATGCTGCTGCGA 9723
4563 TCTGTTAGGCACTGATCATACAGACCTGATGATGCGCAAGAAAGAAAGACTGCTGA 4622
9724 TGATATTTGGGCTAACTGTTGACATTAATGCTGTAAGCGCGCTGATGCTGTA 9783
4623 CGCCATAGCGGCTTGCACGCAACGCAAGGTAACCCACAGCATTAACGCTTTGCTGCA 4682
9784 CGCCATAGCGGCTTGCACGCAACGCAAGGTAACCCACAGCATTAACGCTTTGCTGCA 9843
4683 TTCCGCAACCAACGCTGAAACATTTGAGAACTTTGAAACCATCTGTTGTTTAAAC 4742
9844 TTAGGTCGGCTAATCTGAAACATTCACGAGACCATGATTAATTAATGTTGAAAC 9903
4743 AACCTTTCTCTGGGCAAGTGTGATCTCTGTCGACGCTGTTTATTCTGTTCCGT 4802
9904 AGCGTCTCTCTGGGCTGACGCTGTATACCTCTGCGCTGTCTGCTTAAATGCGCT 9963
4803 GCTTTTCAATGCTGACGCTTTTATTGTTGTTGAGGCGTCTGCTGGGAAAGTGAAG 4862
9964 GTTGCTATGCTGCTGCTCTTTTATTGTTGTTGCGGCGCTTACCTTGGGAAAGTGAAG 10023
4863 CTTTGGAAATGTCGACCACTGTGCGCAATGTTTCGCGGATCCGTAATAGGCTGTG 4922
10024 CTTAAGAACATGTCGACCACTGTGCGCAATGTTTCGCGGATCCGTAATAGGCTGTG 10083
4923 AACGCGAGTTTACCGGCACTTAACCTGAGATACGCTGTCTCATTCGGAATTAAC 4982
10084 AAAAGGACGAGTACCGCGCTCAATTTGGAATTAATCTCATGCTCTCGGAGGTTTGC 10143
4983 CTTCAACTAACAGAGTACGTAACCTGCAATTCACACAGTATTCCTTACACAGAC 5042
10144 CTTTCAACCAACAGAGTACGTAACCTGCAATTCACACAGTATTCCTTACACAGAC 10203
5043 TTAAATGCTGGGCTCCCTGAGTCAAGGCAATCTCAAGGCGGATTAACATCCCG 5102
10204 TCAGATGCTGGGCTCTTGGAAATGTCAGCCGCGCTCAAGCAATATACCTGCAAG 10263
5103 TTTTGGGCTGTATACCTTTCAATGTTGAGGAGGCGCAATGCTTTCTGTGACAGTGA 5162
10264 TCTTTGAGAGGCTGTATACCTTTCAATGTTGAGGAGGCAATATGTTTGTGCAAGTGA 10323

5163 AACACAACTGATGAGGCGTACGTGAGTTGCTCCAGACTGCACTATAGTACAGAG 5222
10324 AACGCATATGATGAGGCGTACGTGAGTTGCTGAGTATGTTGGCGACTACAGCGGC 10383
5223 TCGCACTAAAGTTCAACAGCTGTCTGAAAGTGGCTGCGTATATGATTAACGCAACA 5282
10384 AGGCAATTAAGTGTGATCTCCGCGATGAAGATGAGCTCGTATATGTAACGGGAACA 10443
5283 CCAACCGGCACTGGATGCTGTTGTCATGCGTACAGCGGATTCCTACAGGAGCTGA 5342
10444 CTACAGATTCTTATGATGTGATCGTGAAGCGAGTACACAGGAACGTTAAAGCTGTA 10503
5343 AGTCTATGACAGGCGGATATACGCGGCTTTTACACCTTTGACCATTAAGTGTGATCA 5402
10504 AAGTATATGCTGACCAATTTACAGATTTGTTTACACATTTGATACAAAGGCTGTA 10563
5403 GAAAGGGCTGTTTACACTACGCTTCCCTGATATGAGTATGAAACCAAGAGGCT 5462
10564 ATCCGCGCTGTGTATCACTATGATCTTCCGGAATCGGAGCGATGAACAGAGGCT 10623
5463 TCGGATATTCAGCATCTCCGCTTGAATGCTACAGACATAGTACCGGCTGACATAC 5522
10624 TTGAGATATTCAGCTACCTCTTGAATGCTTACCAAGACCTATGCTGACAGACATTA 10683
5523 GCTGCTGAAGCTTCTGTCAAGAACATCCAGCTCCCTTACACCAAGCAGTATCAGGCT 5582
10684 GCTATCTCAACCTTCCGCGCAAGACGTCATGCTGCTTACACGAGGCGGCTGAT 10743
5583 ATGAATGTGGAAGAACACTACAGACGACCCCTGCAAGAAACAGACCAATTTGATGA 5642
10744 TCGAGATGTGGAAGAACACTACAGGCGCCACATGCAAGAAACGCGCTTTGGGCTGA 10803
5643 AAATTTGAATGAGACCTCTGGAAGCTTAACTGCTTTAGGGGCAATCCCTATCTGA 5702
10804 AGATTTGACGTAATCCGCTTGAAGCGTGAAGCTGCTCAATAGGGAACATTCCTTCTA 10863
5703 TTGACATCTCTGATGACGCTTTGAGATATGATGAAATCAACCAATTTTGAAGTTA 5762
10864 TTGACATCTCTGATGACGCTTTTATCAGAGATGATGATGACCACTGCTTCAACAGTCA 10923
5763 GCTGCAAGTACAGACTGATTAATTTCTGAGACTTTGATGCTTCTTAACTTACAGT 5822
10924 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10983
5823 ACAAAGTACAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5882
10984 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11043
5883 AGGAAGGACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5942
11044 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11103
5943 GCCCAAGCAAAATTTTATGATTTGCTATGCGGCAAGAAAGTCACTGCAATGCTGAT 6002
11104 GCCCAAGGCAAACTTATGATTTGCTATGCTGATGATGATGATGATGATGATGATGAT 11163
6003 GTAAACCAACGCGGACCACTAATTTGAGAACCAATTAAGTGAACCAATTTCCAG 6062
11164 GCAAAACCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11223
6063 CGGAGTTTCAAAACATTTGGAAGCTGTTGCACTGTTTGGGGAGCATCATCC 6122
11224 CCGCATCTCAAAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11283
6123 TCATTTGTTGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6182
11284 TATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11343
6183 GACTGAGGCGGACACTGA-CATAGCGTAAATCTGATGATGATGATGATGATGATGATGAT 6241
11344 GACCGTACGCGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11403
6242 GCATATGCAAGCGCGCT-----TGACACT 6268

[illegible]

RESULT 7

```

US-09-102-248-1
: Sequence 1, Application US/09102248
: Patent No. 6008035
:
: GENERAL INFORMATION:
:
: APPLICANT: Johnston, Robert E.
:
: APPLICANT: Davis, Nancy L.
: APPLICANT: Simpson, Dennis A.
:
: TITLE OF INVENTION: System for the In Vivo Delivery and
: TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
:
: NUMBER OF SEQUENCES: 12
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
: STREET: 1211 East Morehead Street
: CITY: Charlotte
: STATE: No. 6008035th Carolina
: COUNTRY: USA
:
: ZIP: 28234
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/102,248
:
: FILING DATE:
:
: CLASSIFICATION:
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/801,263
:
: FILING DATE: 19-FEB-1997
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Sibley, Kenneth D.
: REGISTRATION NUMBER: 31,665
: REFERENCE/DOCKET NUMBER: 5470-147
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: 919-420-2200
: TELEFAX: 919-881-3175
:
: INFORMATION FOR SEQ ID NO: 1:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 11663 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
: MOLECULE TYPE: cDNA
:
: FEATURE:
:
: NAME/KEY: CDS
: LOCATION: 60..7559
:
: FEATURE:
:
: NAME/KEY: CDS
: LOCATION: 7608..11342
:
: US-09-102-248-1

```

Query Match	36.9%	Score 2396	DB 3	Length 11663
Best Local Similarity	64.6%	Pred. No. 0		
Matches 3751	Conservative 0	Mismatches 1985	Indels 74	Gaps 9
Qy 580	ACTGACGGTATGAAAGCGGAGCCGATATTTTTCATCGGAAACAGGCGCAAGGTCACCTTC			639
Db 5707	ACTCTCTAAACCGGGGTAGGTGGTACATATTTTTCAGACGACACAGGCCCTGGGCACTTTC			5766
Qy 640	AACGAATATCGTACGTCAATGTAACTACAGAACTATATTGGATCGGGCGCTCCATG			699
Db 5767	AAAAGAAATCCGTTCTTCGACAAACAGCTTACAGAACCCGACCTTGGACCGCAATGTTCTGG			5826
Qy 700	AGAAATATTACGCCCGCGCCCTCGATCTCGAAAGAGAAATGTTACAGAAAGAACTCG			759

Db	5827	AAAGATCTACGCCCCGGGTGCTCGACACGTCGAAAGGAAACAAGCTAAACTCAGGTACC	5886
Qy	760	AATTATGCGCCTCTGGAAGAAATAGAAAGCAAGTATCAATGACGAAAAGTAGAAATATGA	819
Db	5887	AGATGATGCCACCGAAGGCCAACAAAAGGACAGTACGAGTCTCGAAAAGTAGAAAACGAGA	5946
Qy	820	AAGCAATTACAGCGGAGGACTCACTTCTGGATTGGGACATATCTATCATCAGAACTGA	879
Db	5947	AAGCCATTAACCACTGAGGAGCTGGCTTTTCAGGGCTACACATGTATMACTCTGCCACAG---	6003
Qy	880	ATCCGTGCAAGTGTACAGAGTCGAATTAATTCCTGTACCAATCTACTCGTCAACGGTAATTA	939
Db	6004	ATCAGCCAGAAATGCTATPAGATCACTTAACCGAAACCATGTATTTCCAGCAGTGTACCAG	6063
Qy	940	ACAGGTTTACATCTGCAGAGAGTCCGGTTAAACGTGCAACTTAAGTTATCCAGAGAAAT	999
Db	6064	CGAACTACTCGACCCCAAGTTGCTGTAGCTGTGTATTAACAATACTATCTGATGAGAAAT	6123
Qy	1000	ACCCTTACGTAGCCAGTTATTTGTAATACAGATGAAATACGATGCGTATCTTGACATGGTGG	1059
Db	6124	ACCGGACGATGCACTTATACATGATCAACCGACGATGCAATGCTTACTTGGATATGTAG	6183
Qy	1060	ACGGCGCATCGTGTCTGATACAGCAACTTTTGTCCGGCTAAACTGAGAACTGAC	1119
Db	6184	ACGGGACGTCGCTTGCTTAATATCTCACTTTTCCCGCCCAACCTTAGAAATTAAC	6243
Qy	1120	CAAGAAGCATAGCTATTTTGCAGCCAGAGATTAAGATAGCGCTCCATCCGCTTATCAGA	1179
Db	6244	CGAAAAGCACAGATATAGAGCCCAAAACATCCCAATGCGTTCATCAGCAATGACAGA	6303
Qy	1180	ATACATTCACAAATGTATTGGCTGCAGCTACTAAAGAAATTGCACGTTACCCAAATGC	1239
Db	6304	ACAGGTTSCAAAACGTGCTCATTTCCCGCATTAAGAAACCTGCACAGTCAACAAATGC	6363
Qy	1240	GAGAAATTAACCTGTCTTAGATTCCGGCGGCACTTAATGTTGATGTTTCAAGAAATAGCAT	1299
Db	6364	GTTGAATGCCCAACTGAGCTACAGCAATTCACGTTGATGCTTTTGAAAATATCAT	6423
Qy	1300	GCAATGATGATGACTCTGGGATACCTTTTGGCGATAACCTTAATCCGCTAACTACAGAGACG	1359
Db	6424	GCAATGACGAGTATTTGGAGGAGATTTCGCCGAAAGCCAAATTAGATCACTACGATGTCG	6483
Qy	1360	TTACGCAATATGTGACAAAGCTGAAGGCGCGGAAGACGACGACATTTGTCGAAATCTC	1419
Db	6484	TTACCGCATATGCTGACGACACTGAAGGCGCTTAAGGCGCGCACTGTTCGCAAAAGACG	6543
Qy	1420	ATTAATCTPAAACCGTTGAGAGATTAACCAATGATTCAAATTCGTATGATGATCTAAAGAG	1479
Db	6544	ATTAATTTGGTCCCATGTGAAAGATGCTTAAGATTAATGCTATGACATGAAAGAG	6603
Qy	1480	ATGTCTAAAGTATCTCCCGGACGAAACATACAGAGAGCGGCTTAAGGTGACGTTATTC	1539
Db	6604	ACGTGAAGTATACACTGTGCAAGAAACACAGAAAGAAAGACGAAAGTACAAATGATAC	6663
Qy	1540	AGGCTGCAGATCCCTCTGTACTACCGCTTACCTTTGCGGAAATCATCGGAAATTAATCCGTA	1599
Db	6664	AAGCGGACGAAGCCCTGGCGACCGCTTAACCTATCGGGAATCAACCGGGAAGTTAGTGGCA	6723
Qy	1600	GACTGATGCGGTGCTTTCGCAATATCCATATCTTCTTGACATGTCAAGCGGAATAT	1659
Db	6724	GAGCTTACGCGGTTTTGTGTACCAACATTCACAGCTCTTTGACATGTCCGCGGAGGACT	6783
Qy	1660	TTGATGGCATTAATTTGCTGAACATTTTCCACCGCGGACCCAGTATTTGAAAACGGACATCG	1719
Db	6784	TTGATGCAATATATGCAAGAACATCTCAAGCAAGGTACCGGTACTGAGACGATATTCG	6843
Qy	1720	CGTCTTTGATAAAGCGAAGACGAGCTATCGCATTTTCGCGCTTGATGATCTCTTAGG	1779
Db	6844	CCTGCTTGCAGAAAAGCCAAAGACGAGCTATAGGGCTTAACCGGCTGATGATCTTTGAGAG	6903
Qy	1780	ACTTAGGTGTCAGCAACCGCTCTTAAATTTGATAGAGCGGCTTGGCAATATCAAT	1839

QY 3963 CCTATCTGAGAGACGTCAGCGAGTGTACTATTAAACACCTTCTGGCAGAGCGCA 4022
Db 9124 CTTATCTGGAGAAATTCATCAGGGAAGTTTACGGAGACCACTCCGGAGAAACATTTA 9183
QY 4023 CTTACGAATGTAGTGTGGACATACAGCACAGGTAATGTGAGACCGGAAAGAAATGA 4082
Db 9184 CGTACAGTGTACAGTGGCGGATTAACAAGACGGAAACGTTACGACCCGTACCGAAATCA 9243
QY 4083 ACGGCTGCATTAAGCAAAACAGTGTGCTTCTCAAGAGGACCAACGAATGGGTCT 4142
Db 9244 CGGGCTGCACCGCATCAAGAGAGTGGCTGTATTAAGAGCAACCAAGAAAGGGGTCT 9303
QY 4143 TCAACTGCGCGGATCTTATTAGGACACAGACCACTGAGCAAGTAATTTGACATTC 4202
Db 9304 TCAACTGCGCGGATCTGATTCAGACAGCGGACCAACGCGGAAATTTGCAATTTGC 9363
QY 4203 CATTCGGTGTACACCGACAGTGTGCGCGGTTCCGTTAGCTACACGCGCTACAGTCA 4262
Db 9364 CTTTCAAGCTGATCCCGAGTACCTGACATGTCCTGTTGGCCACGCGCGCAAGTATAC 9423
QY 4263 AGTGGTTCAAAAGGATCACTCTCACTGATCGCAATCCGACCAATTTGCTGACACGA 4322
Db 9424 ACGGCTTTAAACATACAGCTCTCAATTTGAACAAGAACCATCTGACATCTGTCACAC 9483
QY 4323 GAAATTTGGGGCTCGAGCAGACGCAACAGCAGATGATTAAGAGGTCTACATCCAGGA 4382
Db 9484 GGAATCTAGGGGCAAAACCGGAAACCACTGATGATGATGATGGAACACGGTTAGAA 9543
QY 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTAAGTATGGGATTAACCATGAAACAGTCAAG 4442
Db 9544 ACTTCAACGCTGCAGAGATGGCTGGAAATCATATGGGGCAATCAACGAACAGTAAAGG 9603
QY 4443 TCTGGGCGCAGAGATTCGGACACCAAGCCCAACATGATGGCGGATGATGATCATCTCC 4502
Db 9604 TCTATGCCCAAGATCTGACACAGAGACCTCAACGATGGGCAACCAAAATTAATACAGC 9663
QY 4503 ACTATATCATCGGATCCAGTCTACACTGTCTATGTCGTGTGTCGTCTGTCGCTA 4562
Db 9664 ATTACTATCATCGGATCTCTGTGTACACATCTTAAGCCGTGCGATCAGCTCTGTGGCA 9723
QY 4563 TCTGTAGAGCATGTCATCAGCAGCTTGACATCCCAAGCAAGAAAGAGACTGCTGA 4622
Db 9724 TGATGATGGGTAACGTGTGACAGATTAATGTCCTGTAAGCGCGGCTGAGTGCTGA 9783
QY 4623 CGCCATAGCGCTTGACACCAAGCAAGGTAACCAAGCATTAAGCGGTTTGTGTGCA 4682
Db 9784 CGCCATATGCGCTGGCCCAATGCCGTATTCAACTTCCTGGCACTTTTGTGTGTG 9843
QY 4683 TTGGCCAAACCAAGCTGAACATTTGGAAACCTTGAAACCATCTGTGTGTTAAACACC 4742
Db 9844 TTAAAGTGGGCTAAATGTCTGAACATTCACCGAGACATGATTTATATGTGTGCAACGCC 9903
QY 4743 AACCGTTCTCTGGGACAGTGTGCAATCTCTGGACGCGTTGTTATTTCTGTTCGCT 4802
Db 9904 AGCGTTCTCTGGGTCAGCTGTGTATACCTCTGGCGCTGTGTCTTAATGGCT 9963
QY 4803 GCTTTTATGTCGATGCTTTTATATGTGTGACAGGCTCTGCTGGGAAAGTATACG 4862
Db 9964 GTTCTATGTCGCTGCTTTTATATGTGTGACAGGCTCTGCTGGGAAAGTATACG 10023
QY 4863 CCTTGAACATGCGACCACTGTGCAAAATGTTCCGGGATCCCGTATTAAGCGTGTGCG 4922
Db 10024 CTTAGCAATGTCGACACTTTTCCAAATGTGCAAGATACCGTATTAAGGACATTTGTG 10083
QY 4923 AACGCGAGTTAGCGCGCACTTAACCTGAGATCAGCGTGTCTCATGGAATTAACAC 4982
Db 10084 AAAAGGAGGGGTAGCGCCGCTCAATTTGAGATTAATGTCATGTCCTGAGGTTTGC 10143
QY 4983 CTTCACTTAACAAGAGTACGTGACCTGCAAAATTCACACAGTCAATTTCTTACACACAG 5042
Db 10144 CTTTCAACCAACCAAGAGTACATTTACCTGCAAAATTCACACAGTGTGTCCCTCCCTAAAG 10203
QY 5043 TTAATGTGTGGGGTCCCTGAGTGCAGAGCATCTCAAAAGCGGATTAACATGCGCG 5102

Db 10204 TCAGTGTCTGGGCTCCTTGGAAATGTACGCCGCCGCTCACGCACTATATCTGCAAG 10263
QY 5103 TTTTGGCGGTGTGATCCCTTTTCAATGTGGGAGCGCAATGCTTGTGACAGTGA 5162
Db 10264 TCTTGGAGGGGTGTACCCCTTCAATGTGGGAGAGACAAATGTTTTGGACAGTGA 10323
QY 5163 ACAACAACTGAGGAGGGGTAGGTGAGTTCGCTCCAGCTGACATATAGATCAAGCAG 5222
Db 10324 ACAGCAGATGAGGAGCGGTAGGTGCAATTTGTCAATTTGCGGACATGACGCGC 10383
QY 5223 TCGCACTAAAGTTCAACAGCTGCTGTGAAGTCGGCTGCTGATATGATACGCAACA 5282
Db 10384 AGCGATTAAGTGCATACGCGGCAAGAAAGTAAAGATGCGATATGTATCGGAAACA 10443
QY 5283 CCACCGCGCACTGTGATACCTTTGTCAATGCGCTCACGCCAGGTTCTCACGGGACTGA 5342
Db 10444 CTACCAATTTCTTAAAGTGTAGCTGAACCGAGTCAACACGAGAACGCTTAAAGACTGA 10503
QY 5343 AGGTCATGAGAGGGCGGATATACGCGCTTTTCACTTGAACATTAAGGTGTCATCA 5402
Db 10504 AAGTCATGAGTGAACCAATTTGACATTTGTACACCATTTGATCACAAAGTGTATCA 10563
QY 5403 GAAAGGGCTTTTCAACCTACGACTTCCCTGATGATGAGGCTATGAAACAGAGCGT 5462
Db 10564 ATGCGGCTGTGTGTACATATGATGCTTCCGAAATACGAGGATGAAACAGAGCGT 10623
QY 5463 TCGGCAATTTCAAGCATCTGCTGCTGATGTCTACAGCATATGATACCGGACTGACATAC 5522
Db 10624 TTGAGACATTTCAAGCTACCTCTTGAATGACAAAGCTCATGCGCAGACAGACATTA 10683
QY 5523 GGGTCGTAAGCCTTCTGTCAAGAACATCAAGTCCCTACACCGCAAGCATATCAGGT 5582
Db 10684 GGGTACTCAAGCTTCCGCAAGAACGTGATGTCCCTGACAGCAGGCGCATCTGGAT 10743
QY 5583 ATGAATGTGAAAGAACATCAAGACACCCCTGCAAGAAAGACCAATTTGATGTA 5642
Db 10744 TCGAGATGTGAAAGAACATCAAGGCGCGGCACTGACAGAAACCGCCCTTTGGGTGA 10803
QY 5643 AAATGAATGTGAGACCTCTGCGAGCGCTTAATCTGTGTTACGGGACATCCCTATCTGA 5702
Db 10804 AGATTTGAGTCAATCCGCTTTCGAGCGGTGAGCTGCTATCGGAAATTTCCATTTCTA 10863
QY 5703 TTGACATCCCGATGAGCTTTTGTGATCATCAGAAATCAACCAATTTTGAAGTGA 5762
Db 10864 TTGACATCCCGAAGCTGCTTATATCAGACATCAGATGACACATCTGCTCAACAGTCA 10923
QY 5763 GTCGACAGTGAAGACACTGCAATTTATCTGACAGCTTTGGTGTCTTAACATTCAGT 5822
Db 10924 AATGTATGTCAAGTGAAGTGAATTTATCAGCGGACTTCGAGGAGATGCTAACCTGCA 10983
QY 5823 ACAAGTGAACAGGAGGACATTTGTCAGTTCACTCTCCACTCAACAGACGCTGTTTGA 5882
Db 10984 ATGTATCGACCGGAGGACCAATGCTCTGTAATTCGATTTGAGAGCAGCAACCTCC 11043
QY 5883 AGAAGGACACACATGATGAGTCCGTAGGACAGATTAACATTAATTTAGACATGCA 5942
Db 11044 AAGGTGACAGTTCATGTCCTGAGAAAGAGAGCGGTGACATTTAGACAGCGCA 11103
QY 5943 GCCCACAAGCAATTTATATGTTTCTGATGCGGCAAGATCCACTGCAATGCTGAAT 6002
Db 11104 GCCCAGAGGGAATCTTATGATGCTGTGTGTGAAGAACCAATGCAATGACAGAAAT 11163
QY 6003 GTAAACCAACGCGCGGACCAATATTTGAGAACCAATTAAGTGCACCAAGATTTCCAG 6062
Db 11164 GCAAAACCAACGCTGATCATATGTGAGCAACCCGCAAAAAATGACCAAAATTTCCAG 11223
QY 6063 CGGAGTTTCAAAACAATTTGAAACCTGCTGTGACATGTTTGGGGGAGCATATCC 6122
Db 11224 CCGCATCTCAAAACATTCATGAGTGTGCTGTTTGGCTTTTGGGGGCGCTGCTGC 11283
QY 6123 TCATGTTGTAGACATTAATGATGTGTGCTGACGCTTATGCTTAATTAACACAGTAT 6182

Accession	Sequence	Position
Db	11284 TATTAAATTATAGACCTATGATTTTGGCTTCGAGCAATGATGCTGATACACGCAAGAT	11343
Qy	6183 GACTGAGCGCGGACACCTGA-CATAGCGGTAAACTGATGTACTTCCGAGAAAGCTGTG	6241
Db	11344 GACCGCTACGCCCCAATGACCCGACAGCAAACTGATGTACTTCCGAGAACTGATGT	11403
Qy	6242 GCATTAATGCCACGCGCGCT-----TGACACT	6268
Db	11404 GCATTAATGCCACGCGCTGATATTAAGATCCCGCTTACCGCGGCAATATAGCAACACC	11463
Qy	6269 AAAACTGATGTAATTTCCGAGAAAGCAGTGCATATGCTGTGCAGTGT	6318
Db	11464 AAAACTGAGCTATTTCCGAGAAAGCGCATGATATGCTGCAGTGT	11513

```

US-09-367-764-1
; Sequence 1, Application US/0936764
Patent No. 6583121
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Bell Seltzer Park & Gibson, P.A.
STREET: 121 East Morehead Street
City: Charlotte
STATE: No. 6583121ch Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/367,764
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 60..7559
FEATURE:
NAME/KEY: CDS
LOCATION: 7608..11342
? ? ? ? ?
US-09-367-764-1
Query Match      36.9%; Score 2396; DB 4; length 11663;
Best local Similarity   64.6%; Pred. No. 0;
Matches 3751; Conservative    0; Mismatches 1985; Indels     74; Gaps       9

580 ACTGACGGTATGAAGCGGAAAGGTATTTCTCATCGGAACAAGCCCAAGTAACCCTTC        639
|||||              |||||               |          |||||         |||||

```

Db	5707	ACTGCTTAACCCGGGGTATGGTGGTACATATTTTCGAGGACACAGGCCCTTGGGCACTTGC	5768
Qy	640	AACAGAAATACGATACGTCATATGTAATACTACAGAACTTATTTGGATCGGGCCGTCATG	699
Db	5767	AAAAAGATCGGTTCTCTCAGAAACCAAGCTTACAGAAACCGACTTGGAGCGCAATGTTCTGG	5828
Qy	700	AGAAGATTACGCCCGCCGCTCGATCTCCGAANAGAGAAATTTTACAGAAAGAACTGC	759
Db	5827	AAAGAAATCTACGCCCGCTGCTCCACAGCTCGAAAGAGGAAACACTCAAACTCAGTAC	5886
Qy	760	AATTAATCGCGCTCTGAAGAAATATGAAGCAGGTATCATACGAAAGTAGAAAAATATGA	819
Db	5887	AGATGATGCCCAACCGAAGCCAAACAAAGCAGGTACCGAGTCTCGAAAGATAGAAACCAGA	5946
Qy	820	AAGCAATTAACAGCGGAGCGACTCATTTTCTGATTTGGCACATATCTATCATCAGAGTGA	879
Db	5947	AAGCCATTAAACACTGACGCACTGGCTTTCAAGGGCTACAGACTGTATTAAGCTCTGCACAG---	6003
Qy	880	ATCTCTGCGAGTGTACAGAGTCAATATATCTGTACCAATCTACTCGTCAACGGTAATTA	939
Db	6004	ATACGCGAGATGTCTATMAAGATCACTTACCCGAAACACATGATTAATTCAGCAGTGTACAG	6063
Qy	940	ACAGGTTTAATCATCTGCAGAGGTGCGCGGTTTAAACAGTCAACTTAGTTATCCAAAGAAAT	999
Db	6064	CGAACTACTCTGACCCAAAGTTTGTCTGACTGTTTTHAACATTAATCTGATAGAAATTT	6123
Qy	1000	ACCCTACAGTACGCCAGTATTTATTAACAGATGATACAGTGCATCTTGACATGTGTG	1058
Db	6124	ACCGCAGGTATAGCATCTTATCAAGTACCGACGAGTACGATGCTTACTTGGATATGGTAT	6183
Qy	1060	ACGGCGCATGTGCTGTCTTGATACAGCCACTTTTGTCCGGCTTAACTGAAAGCTTACC	1119
Db	6184	ACGGGACAGTCGCTTGCTTGAATCTGCAACTTTTGTCCCGCCCACTTAGAAGTTAAC	6243
Qy	1120	CAAGAAGCATPAGTATTTTGAGCAGAGATTAAGATCAGCGCGTCCCATTCGCTTACAGA	1179
Db	6244	CGAAABACACGAGTATAGAGCCCCAAACATCCGAGTGGGTTTCAATCAGCATGTACGA	6303
Qy	1180	ATACATTACAAATATGTAATTTGGCTGCAGCTACTAAAGAAATTCGACGTTACCCAAATGC	1239
Db	6304	ACACGTTGCAAAACGTGCTCATTTGCCGCGCATAAAGAACTGCAAGTCACAAAATGC	6353
Qy	1240	GAGAAATTAAGCTGTCTTGATAGATTCGGCGGCATTTTAATGTTGATGTTTCAAGAAATACGAT	1299
Db	6364	GTCGAACGCCAACACTGGACTCAGCGCAATTTCAACGTTGTAATGCTTTCGAAAAATATCAT	6423
Qy	1300	GCAATGATGAGTACTGGGATTAACCTTTCCGATTAACCTTATTCGGCTTAACTACAGAAAC	1359
Db	6424	GCAATGACGATATTTGGAGAGAGTTTCCCGAAAGCCAAATTAAGGATACATCTAATGATTC	6483
Qy	1360	TTACGCAATATGTGACAAACCTGAAGAGGCCGAAGCAGACAGCTTTTGGTSCGAATATCTC	1419
Db	6484	TTTACCGATATCGTGGCCAGACTGAAGGCCCTTAAGGCCGCCGCACTGTTCCCAAGACGC	6543
Qy	1420	ATTAATCTTAAACCGTTGCGAGAGATACCAATGATCAATTCGTGATGATCTTAAAGAGAG	1479
Db	6544	ATTAATTTGTCCTCATTTGCAAGAAAGTGCCTATGATATATTCGTATGACATGAAAGAG	6603
Qy	1480	ATGTCAAAAGTTACTCCCGGACGAAACATACAGAGAGCGGCCCTTAAGTGTGAGTTATTC	1539
Db	6604	ACGTGAAGATTTAACCTCTGGACGAAACACAGAGAAAGAACCGAAAGTACAGATGTATAC	6663
Qy	1540	AGGCTGACAGATCCCTTGTCTACGCTTACCTTTTGGGGATTCATCGGGAATTTATCTGTA	1599
Db	6664	AAAGCCGAGAAACCCCTGGCGACCGCTTACTTATCGGGATTCACCCGGGAGTTAGTGGCA	6723
Qy	1600	GACTGAATGGGCTTCTTCCCAATATTCATACTCTTTCGACATGTCAAGCGGAAATTT	1659
Db	6724	GGCTTAACGCGTTTGGCTTACCAACATTTACACAGGCTCTTTGACATGTGCGCGGAGGACT	6783
Qy	1660	TTTATGTGGAATTTATTTCTGAACATTTTCCACACGCGCACCCAGATTTTGGAAACGACATCG	1719
Db	6784	TTTATGTGAATTCATAGCAAGACACTTCAAGAAAGTGAACCGGATCTGGAACCGATATTCG	6843

QY	1720	CGTGGTTGATAAAGCGAANA	CGACGCTATCGGCATTTGGCGGTGATGATCCTTAGG	1779
Db	6844	CTTGTTGCAAAAAGCCAAAGCA	CGCTATGGCGTTAACCGGCTGATGATCTTTGAAG	6903
QY	1780	ACTTAGGTGTGACCAACCG	CTTTAGTTTGATAGAGCGCGGCTTGGGCATATTCACAT	1839
Db	6904	ACCTGGGTGTGACCAACCA	CTTACGCTTGATCGAGTGGCGCTTTGGAAATATCAT	6963
QY	1840	CTGTGCACCTACCTTACAGGA	CGAGGTTTAAATTGGTGCATGATAATCCGGTATGT	1899
Db	6964	CCACCCATCTGCCACGGGTA	CCCGTTTCAAAATTCGGGCGATATGATAATCCGGAAATGT	7023
QY	1900	TCTTAAGCGTGTGGTCAACA	CACTAGTCAATATCATGATTTGCTAGCAGAGTACTACGTG	1959
Db	7024	TCTTCACGCTCTTTGTCAAC	ACAGTCTTGATATGCGACAGAGTATTGAGAG	7083
QY	1960	AAGCGTTAACACAGTCAG	CGGTGCGGCTTATCGGCGACGATTAACATAGTCATGTG	2019
Db	7084	AGCGGCTTAAACGTC	CAATATGTACAGATTTATCGGCGACGACATTTATACCGAG	7143
QY	2020	TTCGTCTCCGACCTTGA	TGGCGGAGAGATCGGCACTGGCTGAAACATGGAATAAAAA	2079
Db	7144	TAGTATCTGA	AAAGAAATGCTAGAGGTGTGCACCTGGCTCAACATGAGGTATAGA	7203
QY	2080	TTATTTGATGCA	GTATTTGTATCAAGCACCTTACTTGTGGGGGATTTATCCTGGTGG	2139
Db	7204	TCATTTGACGCA	GTATCGGGGAGAGACCACTTACTTCTGGGTGGATTCATCTTTGCAAG	7263
QY	2140	ACCGAGTATACAGGCA	CAACCTCTGAGATGCGAGACCTCTTAAAAAGGCTTTTAAAGCTTG	2199
Db	7264	ATTGCGTTTACTCTCA	ACGCGGTGTGGGGGACCCCTTGAAAAAGCGTGTTTAAATGTTGG	7323
QY	2200	GAAAAACCATTTG	CCAGTCGATATACCAAGACCTCGACCGCGCGGGGCACTGCATATG	2259
Db	7324	GTAACCGGCTCC	CAACGCGAGATGACGAAGACGAAGACGAAGACCGGCTCTGCTAGATG	7383
QY	2260	AAGCAATCCGATGGA	CGAATTGGAATTACGACGAGTTAGTAAGGCGGTAGAAATCCA	2319
Db	7384	AAACAAAGCGCTGG	TTTGTAGATGATATACAGACACCTTTCAGTGGCCGTGCAATCTC	7443
QY	2320	GATTCGAGATCAT	TCTGGCAGGCGCTGATCATCAAGTCTCTGTCCACGTTAGCGGAAAGCG	2379
Db	7444	GGTATGAGATGACA	CAATCACTGCTCTGCTGGCAATTAGAACTTTGGCCAGAGCA	7503
QY	2380	TTAAGAACTTCA	AGAGCATTAAGAGGAGGCCCAATCACCTCTTACGCGCTGACCTTAATAGG	2439
Db	7504	AAAGGCACTTTC	CAAGCCATCAGAGGGGAAATTAAGCATCTCTACGGTGGTCTTAATATGT	7563
QY	2440	TGAGGTAGTAA	-----CAGCACTTACCCACCGCAGATGTTTCCA	2482
Db	7564	CAGCATGATACAT	TTTCATCTGATCTAATACCAACAACACACACATGATATGAGAAATTTCT	7623
QY	2483	TACCTTCAGCTGA	ACTTTTCCACAGATTATACCTTCAAAATCCGATGGCTTACCGAGATCCA	2542
Db	7624	TTAATCATGTCTGG	CGCGCCCTTTCCAGGCCCACTGCGCATGTGAGAGCGCGGGAGAA	7683
QY	2543	AACCTCTTAGCG	CGCTGGAAGCGCTTGGCGCCCGCTGGCTGCTCAATTCGAATGAAAT	2602
Db	7684	GGAGGCAAGCGCG	CCCGCTGATGCTGCGCAATGGGCTGGCTTTCCAAATCCAGCAACTGA	7743
QY	2603	CTTAGAGAGTGA	TAGTCACTTACCTTCAAAAC--GATTCACCTAATCCGCGCGCAG	2660
Db	7744	CCACAGCCGTAG	TGCTTATGTGACAGGCACTAGACCTTCAAAACCCCAAGCCCAAC	7803
QY	2661	GTCCACCGCCAA	AGAAAGAGTGCTCTCTTAAAGCCAAACCTTACTAGCCTTAAAAAGA	2720
Db	7804	GCCGCGCGCG	CGCAGAAAGAGCGGCGCAAAAGAACCAACCAAGCGAAGAAACCA	7863
QY	2721	AGAAAGCAAC	CAAGAGCAAAAC--GCAAGCTTAAACCAAGGAAACGACAACTGA	2777
Db	7864	AAACACAGAGAA	AGAAAGCAACTCTGCAAAACCCGAAAGAGAACGAGCTA	7923

QY	2778	TTGTGTAAGAAATTGAGTACGTGGACAAAGCAATTTCCGATC---	ATGCTGAAAGCGCCAGTGA	2834
Db	7924	TGGCACTTAAATTGGAAGCGCGACAGACTGTTCGACGTCAAAAATGAGAGCGGAGATGTCA		7983
QY	2835	ATGATATATGCGCTGTGTGCGAAGGAAGGCTGATGAAACCACTTCACGTTGAAGAAAAA		2894
Db	7994	TCGGGCAAGCACTGGCCATGGAAGAAAGGTAATGAAACCACTTCAGTBAAGAACTA		8043
QY	2895	TTGATTAATGAGCAATTAGCGGCGGTGAATTTGAAGAGGCTAGCATGTACGACTTGGAGT		2954
Db	8044	TTGACACACCTGTGCTATCAAAAGCTCAAAATTTCACAAAGTCGTAGCATGATGAGAT		8103
QY	2955	ACGGCGACGTTTCCCGAAGACATGAAATCAGACACGCTGCAGTTCACAGGAGCAAAACAC		3014
Db	8104	TCGACAGATTGGCCGCTCAACATGGAAGTAGAGGGTTCACCTTCAACAGTAACACCTTG		8163
QY	3015	CGGGCTTCTCAATGTGGACACACGGCGCGAGTCCAGTATGAAATGGGAGATTTACGTCAC		3074
Db	8164	AAGGTTTCTACACTTGGACCAACGGAGCGGTGCAGTATAGTAGAGGACAGATTTTACATCC		8223
QY	3075	CGAAGAGAGTGGGCGGGAAAGGCCACACGGAAAGCCGATCTGTGACACAGACAGCGCAG		3134
Db	8224	CCCGGGAGTAGAGGAGGACAGAGAAACAGTGGTGTGTCGATTATGATATACTCAGGCCGGG		8283
QY	3135	TTTGTGCTATTGTCTTAGAGGTGCAATATGAGGACACGCTACGGCGCTTTCAGTGTCA		3194
Db	8284	TTGTTCGGATATGTCTCGAAGGGGCTGTATAGGGAAACAAACCGCCCTTTCGGTCTCTCA		8343
QY	3195	CTTGGAAACCGAAAGGGGTGACCTTAAGGGATACCCCGAAGGTTTGAACCGTGT---		3251
Db	8344	CTTGGAAATAGCAAGGGAAGACAAATCAAGACACCCCGAAGGGAACGAAGATGTGTCTG		8403
QY	3252	-----CACTAGTTACAGCGGTATGCGTGTTCGAAATGTCAAGTTCCATGCGACAAAC		3305
Db	8404	CTGACACACGTGTACCGGCATGTGCTTTGTTGAAACGTAGAGCTTCCATGCAATGCGC		8463
QY	3306	CACCCGTGTGCTATTCACTAGCCGACAGACGAACTTCACGTCTTGAAAGAAAGCTCG		3365
Db	8464	CGCCCACTGTCTACCCCGGAACCAATCCAGAGCTCTCGACATCTCTGAAAGAACCTGA		8523
QY	3366	ACAAATCCAAATTAGACACGCTGTGTGAGAAACGCTTGAAAT---		3422
Db	8524	ACCAAGAGGCTTAGACACCTCTGCTCAACGCGATATGTGCGGTGGATGCTCGGCGAA		8583
QY	3423	CCAAACGAGCATTTACCGATGACTTTCACACTGACAGTCCCTTCACTTGGGGTCTGCCGT		3482
Db	8584	GTAAGAAGAGGTCACTGACGACTTTTACCTTTGACACGCCGCTACTTGGGACATGCTCGT		8643
QY	3483	ATTGCAGACACTCAACGCGGTGTTTCAGCCCAATAAAATTGAGACGTGTGGAGCAAT		3542
Db	8644	ACTGTCAACCAATACGAAACCGGTCTTTGACCCGATTAAGATCGAGCAGGTCTGGGATATAG		8703
QY	3543	CTGATGATGATGATGATTAAGATCCAGGATTCGGGCAATTCGGCTAATACAGGACAGCA		3602
Db	8704	CGGACGACACACCATATGCGATACAGACTTCCGCGCATTTGGATTAAGACCAAAGCGAG		8763
QY	3603	CTGGCGGATGTCAACCAAAATTCGTTACATGTCTTTTTCGACCAAGCCATGACATCAAGAG		3662
Db	8764	CAGCAAGCTCAAAATTAATGATCCGCTACATGTGCTCGACAGAGATCAATCTGTAAABAAG		8823
QY	3663	ACAGTATGAGAAATAATAGCTATACGACATCTGAGACCTCGCGCTGTCTTGGCCACAAG		3722
Db	8824	GCACCATGATGATCATCAATCAAGTACGACCTTCAGGACCGGTGAGAAAGCTTAGCTACAAG		8883
QY	3723	GGTACTTCTGTAGCTCAATGTCTCTCAGGTGACATGTAAACGTCAATATCAACGAGC		3782
Db	8884	GATACTTCTCTCTCGCAAGTGTCTTCAGGGGACAGCGTAAACGTTAGCATAGCGAGTA		8943
QY	3783	GAGCATCTGAAATTCATGACCGGTGAGAAAAAGATCAGAGAGAGTTTGTGGTGAAG		3842
Db	8944	GCAACTCAGCAAGCTCATGACAAATGGCCCGCAAGATTAACCAAAATTTGTGTGGACGGG		9003
QY	3843	AGGATACTTGTTCACACCCGTTCATGAAAAGCTGTAAAGTGCACAGTTTACGATCACT		3902

Db 9004 AAAAATATGACCTTACCTCCCGTTACGGTAAGAAATTCCTTGACAGAGTACACCGTC 9063
Qy 3903 TGAAGAGACGTCCTCCGGTACATTAACATGACAGGCGAGCCCAACCGGTATTAAGT 3962
Db 9064 TGAAGAAACAACCCCGGTACATCACTATGACAGGCGGAGCCGACCTGCTTATCAAT 9123
Qy 3963 CCTATCTGGAGAGACGTCAGGCGAAGTATCAATTAACCACTTCTGGCAAGAACGTCA 4022
Db 9124 CCTATCTGGAGAAATCATCAGGAAAGTTTACGGAAGCCACATTCGGGAAGAACATTA 9183
Qy 4023 CCTACGAATGTATGTGCGCACTACAGCAAGGTATCTGTAGCACGCGAAGCAAGATGA 4082
Db 9184 CGTACGAGTGAAGTGGCGCGATTCAGAACCGGAACCGTTACGACCCGTACCGAAATCA 9243
Qy 4083 ACGGCTGACCTAAACCAAAACAGTCATTCCTTAAGAGGCAACCAAGCAAAATGGGTCT 4142
Db 9244 CGGGCTGACCGCCCATCAAGACAGTGCCTTATTAAGCGCAACCAAGAGTGGGTCT 9303
Qy 4143 TCAACTCGCCGGATCTTATTAAGCACACAGCACTCAAGTGAAGTAAATTTGCACTTC 4202
Db 9304 TCAACTCGCCGGATCTGATCAGACACGCGCAACCAAGGCAAAATTTGCACTTC 9363
Qy 4203 CATTCGCTTGAACACGACAGTCTGCGCGTTCCGTTAGCTTCAACGCGCTACAGTCA 4262
Db 9364 CTTTCAAGCTGATCCGAGTACCTGCATGTCTCCGTGTCACGCGCGCAAGTATGAC 9423
Qy 4263 AGTGTTCAGAGGATCAACCTTCACCTGACCTGCAATGTGACCAACATTTGCTGACAGA 4322
Db 9424 ACGGCTTTAAACATCACTACGCTTCAATTAAGACAGACCATCTGACATTTGCTACACCA 9483
Qy 4323 GAAATTTGGGGCTGGAGAGACGCAACGAGAAATGATTAAGAGGTCTACATCCAGA 4382
Db 9484 GAGAGCTAGGGCAAAACCGGAACCAACCTGAATGATCATCGGAACACGCTTAAGA 9543
Qy 4383 ATTTTCTGTGGGGAGAGAGGCTGAGTACGATAGGGTAAACCATGACAGTCAAG 4442
Db 9544 ACTTCACGCTGACCGAGATGCGTGAATATCATATGGGGCAATCAGAACCAATGAGG 9603
Qy 4443 TCTGGGCCCCAGAGTCCGACCAAGGCGAACCCACATGATGCGCGATGATCATATCC 4502
Db 9604 TCTATGCCAAGTCTGACAGAGAACCTTCAAGATGGCCACACGAATTAAGTACAGC 9663
Qy 4503 ACTATTATCATCGGATCCAGTCTACCTGATCTGTGTGTGTGTGTGTGTGTGTGTGCT 4562
Db 9664 ATTACTATCATGCGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 9723
Qy 4563 TCTGTGAGGACCTGATCATCAGCAGCTTGATGCGCAAGCAAGAAAGACTGCTGA 4622
Db 9724 TGATGATTTGGCGTAACTGTGTGACATTAATGTGCGTAAAGCGGCGGTGAGTCTGA 9783
Qy 4623 CGCCATAGCGGCTTGACAGGAAGGATCCACACAGATTAAGCGGTTTTGTGTGCA 4682
Db 9784 CGCCATAGCGGCTTGACAGGAAGGATCCACACAGATTAAGCGGTTTTGTGTGCA 9843
Qy 4683 TTGGGCAACCAAGCGTGAACATTTGAGAACTTTGAACCATCTGTGTGTGAACAC 4742
Db 9844 TTAGTCTGCTATGCTGAACATTTGAACCAAGCAATGATTAATTAAGTGTGAACAC 9903
Qy 4743 AACCTTTCTCTGGCACAGTGTGATCTCTGACAGCGCTTTTATTTCTGTGTGCT 4802
Db 9904 AGCGCTTCTTGGGTCCAGCTGTGTATCTCTGCGCGCTGTGTGTGTGTGTGTGTGTG 9963
Qy 4803 GCTTTTATGCTGATGCTTTTATTTATTTGTGTGACAGCGCTCTGCTGGGGAAGTGAAG 4862
Db 9964 GTTGTCTATGTGCTGCTTTTATTTATTTGTGTGCGCGCTCTACCTGGGAAGTGAAG 10023
Qy 4863 CTTGGAACATGAGCACTGTGCAAAATGTTCCGGGATCCCGTATTAAGCGGTGTGCG 4922
Db 10024 CTTAAGAACATGAGCACTGTGCAAAATGTTCCAGATACCGTATTAAGCACTTGTGTG 10083
Qy 4923 AACGCGAGGTTACCGGCACTTAACCTGAGATCAAGGTCTGTATCTGCAATTAACAC 4982

Db 10084 AAAGGCAAGGTTAGCGCCCGCTCAATTTGAGATTAATCTGTATGTCTCGAGGTTTTC 10143
Qy 4983 CTTCAACTAAGAGATGACGTAACCTGAATTCACACAGTCACTTCCTTACACAGAAC 5042
Db 10144 CTTCAACCAACCAAGATGACATTAACCTGAATTAACCACTGTGTCTCCCTCCCTTAAG 10203
Qy 5043 TTAATGTGCGGGTCCCTGAGTGAAGGATCCCTCAAGCGGATTAACATGCGCG 5102
Db 10204 TCAATGTGCGGGTCTTGAATGTGACCGCGCGCTCAAGCAATATATCTGTGAAG 10263
Qy 5103 TTTTGGCGGTGTATCCCTTTTATGTGGAGGCGCAATGCTTGTGTGACAGTGA 5162
Db 10264 TCTTTGAGGGGTGTATCCCTTCAATGTGGGAGGAGCAATGTTTGGCGACAGTGA 10323
Qy 5163 ACACCAATGAGTGAAGCGTACGTCGATTCGCTCCGACGCACTATATGACAGGAG 5222
Db 10324 ACAGCAATGAGTGAAGCGTACGTCGATTCGATGATGTCGCGACTGACACGCGC 10383
Qy 5223 TCGCACTAAAGTTTACACAGCTGCTGAAGTGGCGCTGCTGATTAATATGCAACA 5282
Db 10384 AGCGATTAAGGTGATCTGCGCGAGTAAGAGTGAAGTGTGTATGTATGCGGAACA 10443
Qy 5283 CCACCGGACCTGATACGTTTGTCAATGAGCGTCAAGCGAGTTCTTACGGGACTGA 5342
Db 10444 CTACAGTTCCTAGATGTAGTGAAGCGAGTCAACAGAACTGTAAAGACTGA 10503
Qy 5343 AGCTATGACAGGCGGATATCAAGCGGTTTTTACCTTTGACATTAAGTGTGATCA 5402
Db 10504 AAGTATGCTGAGCAATTTTACAGATTTTGAACCAATTTGATCAAGGCTGTATCA 10563
Qy 5403 GAAAGGGCTGTTTATCAACATACGACTTCCCTGATATGAGATTAAGAACAGAGGCT 5462
Db 10564 ATCGGCGCTGTGTATCAATGATGATTTCCGGATATGAGAGGATTAAGAACAGAGGCT 10623
Qy 5463 TCGGATATTTCAAGATCTCTGCTTGTATGCTACAGACATATGAGCCGCACTGATAC 5522
Db 10624 TTGAGACATTTCAAGCTACTCTTGACTAAGAAAGCTCATGCGCAGCACAGACTTA 10683
Qy 5523 GGTCTGTAAGCTTCTGTCAAGAACATCCAGTCCCTTAACCAACCAAGATTAAGGCT 5582
Db 10684 GGTACTGTAAGCTTCTGTCAAGAACATCCAGTCCCTTAACCAACCAAGATTAAGGCT 10743
Qy 5583 ATGAATGTGAGAAACAATCAGAGACGACCTGCAAGAAACAGCACTTTGGATGA 5642
Db 10744 TCGAGATGTGAGAAACAATCAGAGCGCCCACTGCAAGAAACCGCTTTGGATGA 10803
Qy 5643 AAATTAAGTGAAGCTCTGCGGTGTAACTGTGTACAGGCAATCTCTATCTGA 5702
Db 10804 AGATTCAGTCAATTCGCTTGAAGGAGTGAAGTGTCTATACAGGAACTATCCATTTCTA 10863
Qy 5703 TTGACATCTGATGACAGCTTTTGTGAGATCATCAAGATCAACCAATTTTAAGATTA 5762
Db 10864 TTGACATCTGATGACAGCTTTTGTGAGATCATCAAGATCAACCAATTTTAAGATTA 10923
Qy 5763 GCTGACAGTGAAGACTGATTTATTTGAGAGCTTTGTGTGTGTGTGTGTGTGTGTGTG 5822
Db 10924 AATGTATGTAGTGAAGTGAAGTGAAGAGGCGGTGACAGTCACTTGTGACACCGGA 10983
Qy 5823 ACAAGCTGACAGGAGGAGATTTGTCCAGTTCACTCCCATCCACAGCAGCTGTTTGA 5882
Db 10984 ATGTATCCGACGGAAGGAGCAATGCGCTGTAACTTGCATTTGAGACAGAACCTTCC 11043
Qy 5883 AGAAGGACCAACATGTGACTGCGTGAAGCAGATTAACCTATTTTGTGACATGA 5942
Db 11044 AAGGTGACAGTTCATGCTCTGAGAAAGAGGCGGTGACAGTCACTTGTGACACCGGA 11103
Qy 5943 GCCCAAGCAAAATTTTATGTTGCTATGCGGCAAGAAATTCACCTGCAATGCTGAAT 6002
Db 11104 GCCCAAGCAAAATTTTATGTTGCTATGCGGCAAGAAATTCACCTGCAATGCTGAAT 11163
Qy 6003 GTAAACCAACCGGCGACCACTAATTTGAGAAACCACTAATGTCGACCAAGATTTCCAG 6062
Db 11164 GCAAAACCAACGCTGATCATATGAGAACACCCGCAAAATAAGACCAATTTCCAG 11223

Db 6883 CATCATTCGACAAAGCCAAAGACGACGCTATGCGCTTAAACCGGTGATGATCTTGAGG 6942
Qy 1780 ACTTGGTGTGACCAACCGCTTATGATTTATAGAGCGCGCTTCCGCAATATCACT 1839
Db 6943 ACTGGGTGTGATCAACCACTACTGATTTGATGAGTGCCTTTGAGAAATATCAT 7002
Qy 1840 CTGTCCACTTACTCAGAGAACGAGGTTAAATTTGTCSCATGATGAAATCCGGTATGT 1899
Db 7003 CCACCATTACTACGGGTACTCGTTTAAATTCGGGCGATGATGAAATCCGAAATGT 7062
Qy 1900 TCTTACGCTGTTTGTCAAACACTAGTCAATATCATGATTCAGACAGTACTACGTG 1959
Db 7063 TCCTCACTTTTGTCAACACAGTTTGAATGTGGTATTCGCCAGACAGTACTAGAG 7122
Qy 1960 AACGTTAAACACGTCAGGTCGGCGCTCTATTCGGCGACATTAACATATGTCATGTG 2019
Db 7123 AGCGGCTTAAACGTCAGATGTGACGCTTCAATGGCGACACAAACATCATATCATGAG 7182
Qy 2020 TCGTCTCCGACACCTTGTATGCGGAGATGCGCCACTTGGCTGAACATGAGTAAAA 2079
Db 7183 TAGTATCTGACAAAGAAATGCTGAGAGGTGCGCCACTCGCTCAACATGAGGTTAA 7242
Qy 2080 TTATGTATGACATTTATGTTATCAAGCAACCTACTTCTGTGGGGAATTTATCTGTGG 2139
Db 7243 TCATCGACGACATCTCGGTGAGAGACCACTTACTTCTGCGCGGATTTATCTTGCAAG 7302
Qy 2140 ACCAATTAACAGGCAACGCTCGACAGTGTGACAGACCTTAAAGGCTTTTAAAGCTTG 2199
Db 7303 ATTCGGTTATCTTCCACAGGTGTCGCGGTGCGACCCCTGAAAGGCTGTGTTAAGTTGG 7362
Qy 2200 GAAACCACTTGGCAGTGTATGATACCAAGACTGGGACCGCGCGGACTGATGATG 2259
Db 7363 GTTAAACCGCTCCACGCGACGACGAGCAAGAAAGCAAGAACGCGCTCTGCTAGATG 7422
Qy 2260 AAGCAATCGATGAAACAGAAATTTGAAATTAACGAGAGTATGAAAGCCGTAGATCA 2319
Db 7423 AAACAAAGCGGTGTTAGATGAGTATTAACAGGCACTTTAGCAGTGGCGTGACGACCC 7482
Qy 2320 GATAGAGATCATCTGCGAGGCTGATATCACTCTCTGTCACGTTAGCCGAAAGCG 2379
Db 7483 GGTATGAGGTAGCAATATTACACCTGTCTCTGCTGCAATTTGAGAACTTTTGGCCAGAGA 7542
Qy 2380 TTAAAGACTTCAAGGACATTAAGAGGAGCCCAATCAACCTTACGGCTGACCTAATAG 2439
Db 7543 AAGAGCAATTCBAGCCATCAAGGAGAAATTAAGCATCTCTACGCTGTCTTAATAGT 7602
Qy 2440 TGACGTAGTAGACAGCACTTAC-----CCACCGGAGAAATGTTCCATACC 2486
Db 7603 CAGCATAGTACATTTCACTGTGACTAATACTAACACCAACCAACATGAAATAGAGATTCT 7662
Qy 2487 CTCAGCTGAATTTTCAACCACTTTTACCTTCAAAATCCGATGCGCTTACCGAGATCCA 2546
Db 7663 TTAAATGCTGCGCGCGCGCTTCCGCGCCCACTGCAATGAGAGCGCGCGAGAA 7722
Qy 2547 CTCCTAGCGCGCGCTGAGGCGCTTTCGCGCCCGCTGCTCTCAATTCGAAGTCTTA 2606
Db 7723 GAGGACAGCGCGCCCGATGCTTCCCGCAATGCGGTGCTTCAAAATCCAGCAATCTGA 7782
Qy 2607 GAGAGTGAATGATCACTTCACTTTCAACACAGTCACTTATCCGCGCGAGGTGCA- 2665
Db 7783 CCACAGCGGTGAGTGCCTTCACTTGAACAGGCACTAGACCTAACCCCACTTCCAC 7842
Qy 2666 -----CCGCGCAAGAAAGAAAGTGTCTTAAAGCAAACTTACTAGCTTAAAGAA 2720
Db 7843 GCCCGCAACCGCGCGAGAGAGAGCGCGCCCAAGCAACCAAGCGAGCGAAAGCA 7902
Qy 2721 AGAAGCAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2777
Db 7903 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7962
Qy 2778 TGTGTATGAGTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2834

Db 7963 TGGCATTAAAGTGGAGCGGACAGATGTTTCAAGTCAAGAACGAGAGAGAGAGAGTCA 8022
Qy 2835 ATGATATGCTGCTGCTTTCGGAGAGAGGCTGATGAAACCACTCCAGTTGAGAGAAA 2894
Db 8023 TCGGCGACGACTGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8082
Qy 2895 TTGATTAATGAGCAATTAAGCGCGCGGAGAAATTTGAAGAGAGAGAGAGAGAGAG 2954
Db 8083 TCGACACCCCTGTGTGTAACAAAGCTCAAAATTTACAAATGCTGACAGATATGAGAGT 8142
Qy 2955 ACAGGAGCTTCCCGAAGCATGAATATGACACAGCTGAGTACACAGCGCAAAACAC 3014
Db 8143 TCGCAGATGTCAGATCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8202
Qy 3015 CGGCTTTTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3074
Db 8203 AAGATTTATTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8262
Qy 3075 CGAGAGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3134
Db 8263 CTCGGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8322
Qy 3135 TTGTGCTATTTGTTTGAAGAGTGCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3194
Db 8323 TTGTGCGATAGTCTCGGTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8382
Qy 3195 CTGGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3251
Db 8383 CTGGAATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGCGCTTTTGAATGTCAAGTTCCATGCGCA 3305
Db 8443 CAGACACCTGCTACGCGCAATGTGTGCTGCGAAATGTAGCTTCCATGCGAGCGCG 8502
Qy 3306 CACCGTGTCTATTAACCTGACGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
Db 8503 CGCCCAATGCTATACCGCGCAACCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8562
Qy 3366 ACAATCCAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
Db 8563 ACCATGAGCGCTAGATACCTGCTCAATGCAATTCGAGTGCAGATGCTGCGCAAA 8622
Qy 3423 CCAAAAGAGATTAACGATGATCACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3482
Db 8623 GCAAAAGAGAGATCACTGACAGATTTTACCTGACAGAGAGAGAGAGAGAGAGAGAG 8682
Qy 3543 CTGATGATGATCGATTTGAATCCAGAGTCTGCGCAATTTGCGCTTCAATAGAGAGAG 3602
Db 8743 CGAGAGATTAACCAATGAGATTAAGAGATTTGCGCGCCAGATTTGAGATTAACCAAGAGAG 8802
Qy 3603 CTGCGAGATGACCAAAATTCGTTTACATGCTTTTCAACAGAGAGAGAGAGAGAGAGAG 3662
Db 8803 CAGCAGGCGCAAAAGATGACCGCTTACATGTGCTTGAACAGAGATTAACCGTTAAAGAG 8862
Qy 3663 ACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3722
Db 8863 GCACCATGATGATCAATGAATTTAGACCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 8922
Qy 3723 GGTACTTCTGTGATGATCAATGTCTCTCAGAGTGAAGTGAACGCTCACTGATACAGAGCG 3782
Db 8923 GATACTTCTCTCTGCAAAATGCGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8982
Qy 3783 GAGCATGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3842
Db 8983 GCAACTGCAACCTGATGATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9042
Qy 3843 AGAGTACTGTTGCTCAACCGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3902
Db 9043 AAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9102

QY 3903 TGAAGAGACGCTCGCCGGGTACATAACATGACAGGCCAGGCCCAACGCCGTATAGT 3962
DB 9103 TGAAGAAACAACCTGACAGGCTACATCACTATGACAGGCCAGGCCGACCTTATACAT 9162
QY 3963 CCTATCTGAGAGAGCGCTCAGGCCAAGTGTACATTTAAACCACTTCTGGCAAGAACGCA 4022
DB 9163 CCTACCTGGAAGAAATCATCAGGGAAGTTTACGAAAGCCGCCACTTGGGAAGAACATTA 9222
QY 4023 CTACGAATGTATGTGGGCACTACAGACAGGTATCTGAGCAACGGAACGAAGATGA 4082
DB 9223 CGTATGAGTGCAGATGGCGGCACATCAAGACCGGAACGTTTCCACCGCCCAAAATCA 9282
QY 4083 ACGGCTGACCTAAAGCAAAACAGTGCATTGCTCAAGAGGCAACAAACGAATGGGCT 4142
DB 9283 CTGATTGACCGCCCATAGACAGTGCCTGCTTATAGAGGACCAACGAAGTGGGCT 9342
QY 4143 TCAACTGCGCGATCTTATTAAGGACAAGACCACTCAGTGCAGATTAATTGACATTTC 4202
DB 9343 TCAACTACCGGACCTTATATGACATGACACACACGCGCCCAAGGAATTTGATTTGC 9402
QY 4203 CATTCGCTTGACACCGACAGTCTGCCCGGTTCCGTTAGCTCACACGCTACAGTCA 4262
DB 9403 CTATTCAAGTTGATCCGAGTACTGACATGCTCCTGTTGCCACGCGCCGAATGTAATAC 9462
QY 4263 AGTGGTTCAAGGATCACCCCTCCACCTGACCTGACATGACCAACCAATTGCTGACACGA 4322
DB 9463 ATGCGTTTAAACATACAGCTCCTCAATTAGATACAGACCACTTGACATTTGCTACACCA 9522
QY 4323 GAAATTTGGGCTCGACAGACGCAACAGACAGATGATTAACAGGCTTACCTCCAGGA 4382
DB 9523 GGAACCTAGGGGCAACCCGGAACCAACACTGATGATGATGGGAAAGCGGTGAA 9582
QY 4383 ATTTTCTGTGGGCGAAGAGGCTGAGTACGTATGGGGTAAACATGAAACAGTCAAG 4442
DB 9583 ACTTCAACGCTGACCGAGATGGCTGGAAATCATATGGGAAATCATAGACCACTGAGGG 9642
QY 4443 TCTGGGCGGAGGATCGGCACAGGCAACCAAGTATGAGGCGCATGAGATCATCATCC 4502
DB 9643 TCTATGCCAAGATCGACACAGAGACCTCTACGAGTGGCCACAGAAATGATACACG 9702
QY 4503 ACTATTAATCATCGGCATCGAGTCTACACTGTCAATGTGCTGTGTGCTGCTTCTTA 4562
DB 9703 ATTAATCAACATCGCATCTGTGTACACATCTAGCCGTGCAATCAGCTACCGTGGGA 9762
QY 4563 TCTGTAGGCACTGACATCATCAGCAGCTTGCCAAAGCAAGAAAGACTGCTGA 4622
DB 9763 TGATGATTTGGGCTAACGCTTGCACTGTATGTGCTGAAAGCCGCGCTGAGCTGGA 9822
QY 4623 CGCATATCGGCTTGACCGAACGCAACGGAACCCACAGATTAGCGGTTTGTGCTGA 4682
DB 9823 CGCATATCGCCTTGCCCAAGCCGCTGAATCCCAACTTGGCTGCACTTGTGCTGCG 9882
QY 4683 TTGCGCAACCAACGCTGAACATTTGGAGAACTTTGAACATCTGTGTTTAAACAC 4742
DB 9883 TTAGGTGGGCAACGTGAACGTTTACCGAGACATGATGATTACTGTGTGCAACGTC 9942
QY 4743 AACGTTTCTGTGGGCAAGTTGCAATCTCTGTGGGAGCGCTTGTATTCTGTTCCGT 4802
DB 9943 AGCGTTTCTGTGGGCTCAGTTGTGCAATCTTTGGCGCTTTTATGATGTTTATGCGCT 10002
QY 4803 GCTTTTATGCTGATCGCTTTTATTTGTTTGAAGCGTGTGCTGGGGAAGGTATACG 4862
DB 10003 GCTGCTCTGTGCTGTCTTTTATTTAGTGTGCGGGGCTTACTTGCGGAAGGTATACG 10062
QY 4863 CCTTGAACATGACACCACTGTGCAAAATGTTTCGGGGATCCCTATTAAGCGTGTGTCG 4922
DB 10063 CTTAGCAATGACCACTGTCTTCCAAATGTGCAACAGATACCTATTAAGCACTTGTG 10122
QY 4923 AACGCGAGTTTACGCGCACTTAACCTGAGATCAGCGTGTCTCATCGGAATTAACAC 4982
DB 10123 AAAAGGAGGGGTATGCCCGCTCAATTGTGAGATCACTGTATCTCTCGGAGGTTTTCG 10182

QY 4983 CTTCAATTAAGAGATAGCTGACCTGCAAAATTTCCACACAGTCACTTCTTCCACACAG 5042
DB 10183 CTTTCAACCAACCAAGATTAATTAATCTGCAAAATTTACACCTGTGTGCTCCCTCCCAAAA 10242
QY 5043 TTAATGCTGCGGGTCCCTGAGTGCAGGCACTCTTCAAGGCGGATTAACATGCGCG 5102
DB 10243 TCAATGCTCGGGCTCTTGTGAATGTGACGCGCGCTCATGACATTAATCTGCAAGG 10302
QY 5103 TTTTGGCGGTGTACCTCTTATGATGGGGAAGCGCAATGCTTCTGTGACATGTAGA 5162
DB 10303 TCTTGGAGGGGTCTACCCCTTTATGTGGGGAAGGCAATGTTTGGACAGTGA 10362
QY 5163 ACACACATGATGAGGCGGTAGCTGAGTTCCTCCAGACTCATAATGATACAGCAG 5222
DB 10363 ACAGCAGATGATGAGGCGGTAGCTGAGTTCCTCCAGACTCATAATGATACAGCAG 5282
QY 5223 TCGCACTAAAAGTTTACACAGCTGTCTGTGAAGTGGCGCTGCTGTATGTATACGCAACA 5282
DB 10423 AGCGATTTAAGTGCACACTGCGCGATGAAGTAAAGTACGCTATGTATGACGGGAACA 10482
QY 5283 CCACCGGCACTGATAGCTTTGTCAATGGCGTCAACGCAAGGTTCTCACGGGACCTGA 5342
DB 10483 CTACCAAGTTTCTTGAATGTGTAGCTGAACGAGTCAACAGGAACGCTTAAAGACTTGA 10542
QY 5343 AGGTATAGCAGGGGCGATATACAGCCGCTTTTCAACCTTTGACATTAAGTGTGCTATCA 5402
DB 10543 AAGTATAGTGCAGACCAATTTTACGATCGTTTACGCAATTCATTAAGTGTGCTATCC 10602
QY 5403 GAAAGGGCTTGTTTTACAACTAGACCTTCCCTGAGTATGAGGCTATGAAACAGAGCGT 5462
DB 10603 ATGCGGCTGTGTGTCACTATGACTTCCGGAATATGAGGAGTAAACAGAGCGT 10662
QY 5463 TCGGCAATTAAGACATCTCTGCTGTGTATGACATAGTATAGCCCGACATGACATAC 5522
DB 10663 TTGAGACATTCAAGCTACCTCTTGAAGAGATCTCATGCGCAGACAGACATTA 10722
QY 5523 GCGTGTGAGACCTTCTGTGAAGAAATCCAGCTCCCTACACCAAGCTATAGGCT 5582
DB 10723 GGTACTGAGCCTTCCGCGCAAGACGTGATGCTCCGTCACAGCGCGCATAGAT 10782
QY 5583 ATGAAATGTGAAGAACCACTCAGAGACCCCTGCAAGAAAGACACCAATTTGATGTA 5642
DB 10783 TTGAGATGTGAAGAAACCACTCAGAGCGCGCCATGCAAGAAACCGCACTTTGGGGTGA 10842
QY 5643 AAATTAAGTGAAGCTCTGCGAGCGCTTATCTGTGTTACGGGCAATCCCTATCTGGA 5702
DB 10843 AGATTAAGTGAATTCGCTCGAGCGGTGACGTGTATAGGGGAACATTTCCATTTCTA 10902
QY 5703 TTGACATCCCTGATGACGCTTTTGTGAGATCATCAGATCAACCAACAAATTTAGAACTTA 5762
DB 10903 TTGACATCCCGAAGCGTCTTTATGAGACATCAGATGACCACTGTGTCTCAACAGTCA 10962
QY 5763 GCTGCAAGTGAAGACTGATTTATTTGACAGACTTTGTGTGTTCTTAACATTAACAT 5822
DB 10963 AATGTAAGTCAAGTGAAGTCACTTATTAAGCAACTTCCGCGGAGTGGCCACCTGAGT 11022
QY 5823 ACAAGCTGACAGGAGGACATTTGTCAGTTCATCTCCCTCCACAGACGCTGTTTGA 5882
DB 11023 ATGATTCGACCGCGAAGGTCAATGCCCTGATATTTGACAGCAAGCAACTTCC 11082
QY 5883 AGGAAGGACACACATGTGATGCGGTAGGAGCACTTAACATCACTAATTTAGACATGCA 5942
DB 11083 AAGATGCAAGTACATGTCTTGAGAAAGAGAGCGGTGACATGACATTTAGACCGCGCA 11142
QY 5943 GCCCACAAGCAAAATTTATAGTTTGTGCTATGCGCAAGAAAGTCCACCTGCAATGCTGAAT 6002
DB 11143 GTCCACAGGGAACCTTATGATGCTGTGTGGAAGAAACAAATGCAATGCAAGAAAT 11202
QY 6003 GTAAACCAACGCGCGACACATTAATTTGAGAACCAATTAAGTGCACAAAGATTTCCAG 6062
DB 11203 GTAAACCAACGCTGACATATGTGAGACACCCCGACAAAAATTAAGAAATTTCCAG 11262
QY 6063 CGGAGTTTCAAAACATCTTGAAGTGGCTGTGATGCTGTTTGGGGAGCATCATCC 6122

Db 11263 CCGCATCTCAAAAACATCATGAGTGGCTCTTGGCCCTTTCCGCGCGCTCGTGC 11322
Qy 6123 TCATGTGTGAGGACTTATAGTGTGCTGAGCTCTATGCTTAAACACAGTAT 6182
Db 11333 TATTATATATAGGACTTATAGTGTGCTGAGCTCTATGCTTAAACACAGTAT 11382
Qy 6183 GACTGAGCGGAGCACTGA-CATAGCGTAAACTGATGCTTCCAGAGAGCGTGT 6241
Db 11383 GACCGCTACGCCCAATATATATGCTGAGCTCTGCTGAGCACTGAAT 11442
Qy 6242 GCATATGCGACGCGCGCT 6261
Db 11443 GCATATGCTATGAGCTGT 11462

RESULT 10
US-09-102-248-8
; Sequence 8, Application US/09102248
; Patent No. 6008035
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the in vivo delivery and
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 6008035th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,248
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/801,263
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11703 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-102-248-8

Query Match 36.8%; Score 2388.4; DB 3; Length 11703;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3698; Conservative 0; Mismatches 1981; Indels 41; Gaps 8;

Qy 580 ACTGACGTATGAACCGGAGCTATTTCTCATCGAAACAGCCAGGCTCACTTC 639
Db 5746 ACTGACGTATGAACCGGAGCTATTTCTCATCGAAACAGCCAGGCTCACTTC 5805
Qy 640 AACGAAATTCAGTACGTCAATGTAACAGAACTATATATGATGGGCGCTCATG 699
Db 5806 AAAAGAGTCCGTTCTGAGAACCAAGCTTACAGAACCAAGCTTGGAGCGCAATGCTCG 5865

Qy 700 AGAATATTACGCCCCCGCTCGATCTGAAAGAGAGAAATGTACAGAGAACTGC 759
Db 5866 AAGAAATTCATGCCCCCGCTCGATCTGAAAGAGAGAAATGTACAGAGAACTGC 5925
Qy 760 AATTATGCGCTTGAAGAAATGAAGCAGTATCAATCAAGAAAGTAAATATGA 819
Db 5926 AGATGATGCCACCGAAGCCAAAGAAAGTATGACAGCTCTGTAAGTAAATATGA 5985
Qy 820 AAGCAATTCAGCGGAGGAGTATCTTGTGATTTGGGCACTATCTATCATCAGAGTGA 879
Db 5986 AAGCAATTCAGCGGAGGAGTATCTTGTGATTTGGGCACTATCTATCATCAGAGTGA 879
Qy 880 ATCTGTGAGTGTATACAGATGATCAATTCCTGTACCAATCTACTCTGCAACGGTAATTA 939
Db 6043 ATCAAGCCAGATGCTATTAAGATCACTATCCGAAACCTTATCTCCAGTACCGG 6102
Qy 940 ACAGGTTTACATCTGACAGAGTCCGCTTAAACGTCMACTTATGTTATCCAGAGAAAT 999
Db 6103 CGAATCTCTCGATCCAGATTCGCTGTAGCTGTGTATCAATCTATCTGATGAGAACT 6162
Qy 1000 ACCCTACAGTACGAGTATTTATTAACAGATGAATGATGCTATCTTACATGCTG 1059
Db 6163 ATCCGACAGTACATCTTATCAGATTACTGACGATGATGCTTACTTGTATGATG 6222
Qy 1060 ACGGCGATCGTGTCTAGATACAGCACTTTTGGCGGCTAAACTGAGAACTAC 1119
Db 6223 ACGGCGATCGTGTCTAGATACAGCACTTTTGGCGGCTAAACTGAGAACTAC 6282
Qy 1120 CAAGAAAGCATAGCTATTTTGCAGCAGAGATGATGATGATGATGATGATGATGATGAT 1179
Db 6283 CAAGAAAGCATAGCTATTTTGCAGCAGAGATGATGATGATGATGATGATGATGATGAT 6342
Qy 1180 ATACATTTCAAAATGATTTGCTGCTGAGTACTTAAAGAAATGCAAGTTTCCAAATGC 1239
Db 6343 ACAGCTACAAATGATTTGCTGCTGAGTACTTAAAGAAATGCAAGTTTCCAAATGC 6402
Qy 1240 GAGAAATACCGTCTTATGATTTGCTGCTGAGTACTTAAAGAAATGCAAGTTTCCAAATGC 1299
Db 6403 GAGAAATACCGTCTTATGATTTGCTGCTGAGTACTTAAAGAAATGCAAGTTTCCAAATGC 6462
Qy 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
Db 6463 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6522
Qy 1360 TTACGCAATATGTAACAAGGTGAAGGCGGAAAGCAGCAATGTTTCCGATATCTC 1419
Db 6523 TCACCGCAATATGTAACAAGGTGAAGGCGGCGGCACTATTTTCAAAAGACGT 6582
Qy 1420 ATATCTTAAACCGTCTGAGAGATACCAATGATGATGATGATGATGATGATGATGATGAT 1479
Db 6583 ATATCTTAAACCGTCTGAGAGATACCAATGATGATGATGATGATGATGATGATGATGAT 6642
Qy 1480 ATGTCAAGTATCTCCGCGCAGAAACATACAGAGAGCGGCTTAAAGTGCAGGTTATTC 1539
Db 6643 ACGTGAAGTATCTCCGCGCAGAAACATACAGAGAGCGGCTTAAAGTGCAGGTTATTC 6702
Qy 1540 AGGCTGAGATCCCTTGTCTACGCTTACCTTTGCGGAGTCAATGCGGATTTAGTCCGTA 1599
Db 6703 AAGCGCAGAAACCTTGTCTACGCTTACCTTTGCGGAGTCAATGCGGATTTAGTCCGTA 6752
Qy 1600 GACTGATGCGGCTCTTCTGCAATATCAATCTCTTCTGCAATATGCGGAGAAATTT 1659
Db 6763 GACTGATGCGGCTCTTCTGCAATATCAATCTCTTCTGCAATATGCGGAGAAATTT 6822
Qy 1660 TTGATGCGATTTATGTAACATTTTCAACAGCGGCGAGTATGTAAGAAAGCAATTCG 1719
Db 6823 TTGATGCGATTTATGTAACATTTTCAACAGCGGCGAGTATGTAAGAAAGCAATTCG 6882
Qy 1720 GTCGTTGATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1779
Db 6883 CATCATTCGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6942

1780 ACTTAGGTGCGACCAACCGCTCTTAGATTGATAGAGCGCGCTTGCGCAATATACAT 1839
1843 ACCGTGGTGTGATGACCACTACTGCTGATGAGTGGCTTTGGAGAAATATCAT 7002
1840 CTGTGACCTTACCTGACGAGGAGGTTTAAATTGGTCCATGATGAAATCCGATATG 1899
7003 CCACCCATCTACTACGGGTACTGTTTTAAATTGGGGGCGATGATGAAATCCGAAATGT 7062
1900 TCTTAAAGCTGTGTGTGACCACTAGTCAATATCATGATTGCTAGCAGAGTACTAGTGT 1959
7063 TCCCAACACTTTTGTGCAACACAGTTTGAATGTGTTATGCGCAGCAGAGTACTTAAAG 7122
1960 AACGGTTAACACCTGACGCTGCGCGCTCTATCGCGCAGATTAACATAGTGCATGTG 2019
7123 AGCGGCTTAAACCTCCAGATGTGTGACGCTTATGGCGACGACAACTCATATGAGAG 7182
2020 TCGTCTCCGACACCTTGAATGGCGAGAGATGCGCACTTGGCTGAACATGGAATGAAAA 2079
7183 TAGTATCTGACAAAGAAATGCTGAGAGTGCAGCCTGCTCAACATGAGGTTAAGA 7242
2080 TTATTGATGCGATTATTGATCAAAAGACCTTCTGTGGGGGATTTATCCTGTGG 2139
7243 TCATCGACGAGTATCGGTAGAGACCACTTCTGTGGCGGATTTATCTTGAAG 7302
2140 ACCAGATTAACAGGACGCTGACAGATGCGAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
7303 ATTGGTTACTTCCACAGCTGCGCGCTGGCGAGACCTTAAAGGCTTTTAAAGTTGG 7362
2200 GAAACCATTTGCCATGATGATATCCAAAGCTGCGACCGCGCGGCACTGATGATG 2259
7363 GTAAACCGCTCCAGCGCAGACGAGCAAGAGCAAGAAAGCGCCTCTGCTAGATG 7422
2260 AAGCAATGCCATGGAACAGATTTGGAATTAGCAGAGTGTGAAGGCGCTAGAAATCCA 2319
7423 AAAAAGGCGGTGTTTGAATAGATGATTAACAGGCACTTGAAGTGGCTGTGACGCC 7482
2320 GATACGAGATCATCTGCGAGCGCTGATCATCGTCTGTCTGACAGTTAGCCGAAAGCG 2379
7483 GGTATGAGTGAACATATTACACTGTCTCTAGTGGCATTTAGAACTTTTCCAGAGCA 7542
2380 TTAAGAACTTCAAGAGCATTAAGAGGAGCCCAATCACTCTGACGCTGACCTTAATAG 2439
7543 AAAGGCACTTCCAGGCTCAGAGGAGGAAATAAAGCATCTCTACGATGCTCTAAATAGT 7602
2440 TGAGTGTATGACAGGACCTTAC-----CAACGCGCAGAAATGTTTCAATACC 2466
7603 CAGATATGATATTTCACTGACTAATTAATACTAACACACACACATGATAGAGATTTCT 7662
2487 CTCAGCTGAACCTTTCACACAGTTTACCTTACAAATCCGATGAGCTTACCGAGATCAAA 2546
7663 TTTAACTGTCTGGCGCGCGCTTCCGCGCTTCCGCGCTGCTGATGTGAGCGCGGAGAA 7722
2547 CTCCTAGCGCGCGCTGAGGCGCTTTCGAGCCCGCTGCTGCTCAAAATGCAAGATCTTA 2606
7723 GAGGAGGCGCGCGCGCTGCTGCGCAACGAGGCTGCTTCAAAATCCAGCAACTGA 7782
2607 GGAGGTGATGTGATCACTTGAATTTCAAAACAGATCACTTAATTCGCGCGCAGGTCCA- 2665
7783 CCACAGCGGTAGTGTCTGATTTGAGACAGGCAACTGACCTTCAACCCCGACGTCCAC 7842
2666 -----CCGCAAAAGAAAGAGTGTCTTAAAGCCAAACCTTACAGCTTAAAGAA 2720
7843 GCCCGCCACCGCGCGCAGAAAGAGCGCGCGCCAGCAACACGAAAGCGAAAGAAACCA 7902
2721 AGAAGCAGCAAGCCAAAGAGAGCAAAAC---GCAGGCTTAAACAGAGAAACGACAGCTA 2777
7903 AAACGACAGAAAGAAAGAAAGCACTGCAAAACCAAAACCGGAAAGAGACAGCGCA 7962
2778 TGTGTATGAAGTGTGAGTGTGAGCAAGACATTTCCGATCATG---CTGAACGCGCAAGTGA 2834
7963 TGGCACTTAAGTGTGAGGCGCAGCATTTGTTGAGTCAAGAAACGAGGACGAGATGTGCA 8022
2835 ATGATATGCTGCGCTTGTGGAAGAGAGGCTGATGAACCACTCAAGTTGAAGAAAA 2894

8023 TCGGCGACGACATGGCCATGAGAGAAAGGTATGAAACCTCTGACAGTGAAGAACCA 8082
2895 TTGATTAATGACATTAAGCGCGCTGAAATTTGAAGAAAGCTAGCATGATGAGTGTGAG 2954
8083 TCGACCAACCTGTGTATCAAAAGCTCAAAATTTTCCAAAGTGTGACGATACGATGAGT 8142
2955 ACGGCGAGTGTCCCGAAGCATGAATTCAGACACGCTGCGAGTACACAGCGCAAAACCA 3014
8143 TCGCAGCTTCCAGCTACATGAGAAAGTGAAGCATTTACCTTACACAGTGAACACCCCG 8202
3015 CGGCGCTTCAACACTGACCAACAGCGCGCACTCCAGTATGAGAAATGGAATTTTACCGTAC 3074
8203 AAGGATTTATTAATGAGCAACAGAGCGGTGAGTATGAGAGGATTTTAAACATCC 8262
3075 CGAGAGAGTGTGGCGGGAAGGCGACAGCGGAAGACCGATCTCTGACCAACAGAGCGAG 3134
8263 CTCGCGAGTGAAGAGCGAGAGAGACAGCGGTGTCTGATCAATGATTACTCCGCTCGGG 8322
3135 TTGTGGCTATGTTCTTAGAGAGTGCAAATGAGGGCAAGCGTACGCGGCTTTCAGTGTCA 3194
8323 TTGTGCGATGATCTCTGAGAGCTGATGAGAAACAGAACTGCTTTCGCTGTCA 8382
3195 CTGGAACCAAGAAAGGAGTACCATTAAGGATATCCCGAAGGTTCTGAACCGTGT--- 3251
8383 CTGGAATGATTAAGGAAAGACATTAAGACACCCCGAAGGGAAGAAAGATGTCG 8442
3252 -----CACTAGTTACAGCGCTATGCGTGTCTTTCGAATGTCACTTCCATGGAACAA 3305
8443 CAGCACACCTGCTGACGCGCAATGTGTTGCTCGAAATGTGACCTTCCATGGAACCGCC 8502
3306 CACCGGTGCTATTAACCTGACCGCAAGCACTGACAGTGTCTGGAAGAAAGCGTGC 3365
8503 CGCCCATCTCTATACCCCGGAACCTTCCAGAGCTTCCGATCTCTTGAAGAAAGCTGA 8562
3366 ACAATCCAAATTAACAGACGCTGAGAGACGCTTTGAATGTCATCAC---GCCGAC 3422
8563 ACCATGAGGCTTACGATACCTGCTCAATGCAATTTGGCGGTGCGATGCTGCGAGAA 8622
3423 CCAAGCAGACATTAACGATGATCTTCACTGACAGTCCCTTACCTGCGGCTTCTGCGCT 3482
8623 GCAAAAGAACGTCACAGGACCTTAACTTGAACAGCCCTTCTTGGGACATGCTGCT 8682
3483 ATTGACAGACCTGACGCGGTGTTGAGCCCAATTAATAATGGAAGCTGTGGAACGAAT 3542
8683 ACTGCCACCTATGAAACGCTGCTTACGCTTGAAGTGAAGCAGCTTGGGACGAAG 8742
3543 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3602
8743 CGGACGATTAACCATACGATACGATACGATACGATACGATACGATACGATACGATAC 8802
3603 CTGCGGATGATCAACAAATTTCCGTTTACATGCTTTTCAACAGCAGCATGATCAAGAA 3662
8803 CAGCAACGCGAAACAGTATACGCTACATGTGCTTGAACAGGATTCACACGTTAAAGAA 8862
3663 ACAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3722
8863 GCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8922
3723 GGTATCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3782
8923 GATATCTTCTCTCTGCAAAATGCTTCCAGGAGCAGCGTATGAGTATGATGATGATGAT 8982
3783 GAGCATTTGGAATTTATGACAGCTGTGAGAAAGAAAGATCAGAGAAATTTGTGCTGAT 3842
8983 GCAATCAGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9042
3843 AGAGTACTTGTGTTCCACCGCTCATGAGAAAGTGTGAAGTGCACGTTTACGATGAT 3902
9043 AAAAATATGATCTACCTCCGCTTACGCTTAAATAATTTCTTGCACAGTGTATGACCTGC 9102
3903 TGAAGAGAGCTGCTGCGGTATATACATGACAGGCGCAGCGCTATTAAGT 3962

Db 9103 TGAAGAAACAAGCTGAGCTATCATCTATGCAAGGCCGGGACCGCACGCTTATACAT 9162
Qy 3963 CCTATCTGGAGAGCGTCAGGCGAAGTGTATTAACACACCTTCTGGCAAGAGCTCA 4022
Db 9163 CCTACTGGAAAGATCATCTAGGAAAGTTACGCAAGCCGCGCATCTGGAAAGAACATTA 9222
Qy 4023 CCTACGAATGTATAGTGGCGACTACAGCAAGATCTGTAGCAAGCGCAAGAAATGA 4082
Db 9223 CCTATAGAGTCAAGTGGCGGACTACAGACCGGAAACGTTTCGACCGCGACCGAATCA 9282
Qy 4083 ACGGCTGCATTAAGCAAAACAGTCACTTCTCAAGAGCGACCAAGAAATGGGTCT 4142
Db 9283 CTGGTTGACCGCCCATCAAGCAGTGTGTCTTATAGAGCGACCAAGAAAGTGGGTCT 9342
Qy 4143 TCAACTCGCGGATCTTATTAGGACACAGACCACTGAGTGAAGTAAATGGCACTTC 4202
Db 9343 TCAACTCACCGGACTTGTATCAACATGACAGCACACGCGCCAGGAAATTTGCACTTTC 9402
Qy 4203 CATTCGCTTGAACCGCAGTCTGCGCGTTCCTTATGCTCAACGCTTACAGTACGA 4262
Db 9403 CTTTCAAGTTGATCCGAGTACCTGATGTCTCTGTTTCCACCGCGCGAATGTATAC 9462
Qy 4263 AGTGTTCAAAGGATCACTCTTCACTGACTGCAATGCGCAACATTGCTGACAAAGA 4322
Db 9463 ATGGCTTTAAACACATCACTCTTCAATTGATACAGACCACTTGCATTGCTCACACCA 9522
Qy 4323 GAAATTTGGGGCTGGAGAGCAGCAACGACAGATGATTAAGGTCTATATCCAGA 4382
Db 9523 GGAAGCTAGGGGCAAAACCGGAACCAACATGAAATGATCTCGAAGAAAGGTACAGA 9582
Qy 4383 ATTTTCTGTGGGGAGAGAGGCTGTGATAGTATGGGTTAACATGAAACAGTACAG 4442
Db 9583 ACTTACCGTGCACCGAGATGCGCTGTGATATCATATGGGAATATCATAGCCAGTGAAG 9642
Qy 4443 TCTGGGCCAGAGATGCGGACCAAGGCGACCAATGATGCGCGATGATCATATCC 4502
Db 9643 TCTATGCCAAGATGACGACGAGGACCTTCAAGATGCGCACAGAAATGATACAGC 9702
Qy 4503 ACTATATCATGAGGATCCAGTCACTGATATGTGTGTGTGTGTGTGTGTGTGTGT 4562
Db 9703 ATTACTTACCATGCACTCTGTGTACCATCTTATGCGGTGCGATCAGTACCGTGGCA 9762
Qy 4563 TCTCTGTAGGACCTCATCATCAGAGCTTGTGATGCGCAAAAGCAAGAGACTCTGA 4622
Db 9763 TGATATTTGGCGTAAACGTTGAGTATATGCTGTAAACGCGCGCTGATGCTCTGA 9822
Qy 4623 CGCCATAGCGCTTGCACCGAAAGCAAGGATCCACAGCATTAAGCGGTTTGTGTGCA 4682
Db 9823 CGCCATAGCGCTTGCACCGAAAGCGGATCCCAACTTCTGTGCACTCTGTGTGCG 9882
Qy 4683 TTGCGCAACCAACGCTGAATAATTGGAGAACTTTGAACCATCTGTGTGTTAAACAC 4742
Db 9883 TTAGTGTGGCCATCTGAAGATTCACGAGACATGATTAATTTGTGTGCAAGTCC 9942
Qy 4743 AACCGTTTCTGTGGGACAGTTGTGATCTCTGTGACGCGCTTGTATATTTGTCGCT 4802
Db 9943 AGCCGTTTCTGTGGGACAGTTGTGATCTCTGTGACGCGCTTGTATATTTGTCGCT 10002
Qy 4803 GCTTTTCACTGTGATGCTTTTATTTATTTGTGACAGCGCTCTGCTGTGGGAGGTAGA 4862
Db 10003 GCTGTCTGTGCTGCTGCTTTTATTTATTTGTGCGCGGCTTACCTGTGGGAGGTAGA 10062
Qy 4863 CCTTGAACATGCGACACTGTGCGAAATTTTCCGGGATCTCCGTTATAGGCTGTGTG 4922
Db 10063 CTTAAGAACATGCGACACTGTGCGAAATTTTCCAGATACCGTATAGGCACTTGTGTG 10122
Qy 4923 AACGCGAGGTTTACCGGCACTTAACTGTGATACGCGTCTCATGCGAATTAACAC 4982
Db 10123 AAAGGCGAGGTTATCCCGCTCAATTGTGATATCATGTCTGTCTGTGAGGTTTTC 10182
Qy 4983 CTTCAACTTAACAGAGTATGATCTGTGAAATTTCAACAGTCAATTTCTTACCAAG 5042
Db 10183 CTTTCAACCAACCAAGATATTAATCTGTGAATTTCAACAGTGTGTGCTTGTGCTTGTG 10242

Qy 5043 TTAATGTGCGGGTCCCTCGAGTGCAGGCACTCTCAAGGCGGATTAACATGCGCG 5102
Db 10243 TCAATGTGCGGGTCCCTCGAGTGCAGGCACTCTCAAGGCGGATTAACATGCGCG 10302
Qy 5103 TTTTGGGCTGTATCCCTTTATGTTGGGAGGCGCAATGCTTCTGTGACAGTGA 5162
Db 10303 TCTTGGAGGGGTCTAACCTTTATGTTGGGAGGCGCAATGTTTTCGACAGTGA 10362
Qy 5163 ACACACATGATGAGGCGTATGCTGAGTGTCTTCCAGACTGCACTATGATCCGAG 5222
Db 10363 ACAGCAGATGATGAGGCGTATGCTGAGTGTCTTCCAGACTGCACTATGATCCGAG 10422
Qy 5223 TCGACATAAAGTTTACACAGCTGTCTGAAGTGTGGCTGTGATTAACGCAAC 5282
Db 10423 AGGCAATTAAGTGCACCTGCGGATGAAGTGAAGTGTGATTAAGTGAAGTGA 10482
Qy 5283 CCACCGCACCTGTGATGCTTGTCAATGCGCTGACCGCAGGTTCTTACGCGGACTGA 5342
Db 10483 CTACAGTTTCTGATGCTGATGCTGAAGGAGTGCACAGGAACTGTAAAGCTTGA 10542
Qy 5343 AGTCAATGCAAGGCGCATATCAGCGCTTTTCACTTGTGACATTAAGTGTGATCA 5402
Db 10543 AAGTCAATGCTGCAACCAATTTACAGCATGTTTACGCGCATGATCAATAGTGTATCC 10602
Qy 5403 GAAAGGGGCTTTTACACACTACGCTTCCCTGAGTATGAGGCTATGAACAGAGG 5462
Db 10603 ATCGGCGCTGTGTACACTATGCTTCCGAAATGAGGCGATGAACAGAGGCT 10662
Qy 5463 TCGGCAATTTCAAGCATCTCTGCTTGAATGCTACAGCATATGATGCGCGCATATAC 5522
Db 10663 TTGAGACATTTCAAGTATCTCTTGAATGCTACAGCATATGATGCGCGCATATAC 10722
Qy 5523 GCGTGTGAGCTTCTGTCAAGAACTCAAGTCTTCAACCCAGCACTATGAGGCT 5582
Db 10723 GCGTGTGAGCTTCTGTCAAGAACTCAAGTCTTCAACCCAGCACTATGAGGCT 10782
Qy 5583 ATGAATGTGAGAAACATCTAGGAGCGCCCGGCAAGAAACAGCACTTTGTGATGA 5642
Db 10783 TTGATGTGAGAAACATCTAGGAGCGCCCGGCAAGAAACAGCACTTTGTGATGA 10842
Qy 5643 AAATTTGAGTGAAGCTCTGCGAGGCTTAACTGTGCTTAAAGGCACTATCTGCA 5702
Db 10843 AGATTGCAATTAATCCGCTCGAGGCTGAGCTGTATTAAGGAACTTTCCATTTCTA 10962
Qy 5703 TTGACATCTGTATGAGCTTTTGTGATCATGAATCAACCAATTTTGAAGTTA 5762
Db 10963 TTGACATCTGTATGAGCTTTTGTGATCATGAATCAACCAATTTTGAAGTTA 10962
Qy 5763 GCTGCAAGTGAAGCTGATTTATTTCTGAGACTTTGTGATCTTAACTATGAGT 5822
Db 10963 AATGTAAGTGAAGTGAAGCTTATTTCAAGCACTTGTGAGGAGTGGCCACCTGTGAGT 11022
Qy 5823 ACAAGTGAAGGAGGAGCATTTGTGAGTCACTTCCATCCAGCAGCTGTTTGA 5882
Db 11023 ATGATGTGAGGAGGAGCATTTGTGAGTCACTTCCATCCAGCAGCTGTTTGA 11082
Qy 5883 AGAAGCGACCAACATGTAAGTCCGTAGGCAACATTAACATTTTGAACATGA 5942
Db 11083 AAGAGTGAAGTGAAGTCACTTGTGAGAAAGAGGCTGACATTTTGAACCGGCA 11142
Qy 5943 GCCCAAGCAAAATTTATTTATTTGCTATGCGGCAAGTATCCACCTGCAATGTAAT 6002
Db 11143 GTTCAAGCGCAACTTATGCTATGCTGTGTGAGAAAGCAACATGCAATGCAAT 11202
Qy 6003 GTTAAACCAACGCGGAGCACTAATTTGAGAAACCAATTAAGTGAACCAATTTCCAG 6062
Db 11203 GTTAAACCAACGCGGAGCACTAATTTGAGAAACCAATTAAGTGAACCAATTTCCAG 11262
Qy 6063 CGGAGTTTCCAAACATCTTTGAAGTGTGCTTGTGCACTGTTTGGGGAGCATATCC 6122
Db 11263 CGGAGTTTCCAAACATCTTTGAAGTGTGCTTGTGCACTGTTTGGGGAGCATATCC 11322

Oy	6123	TCATTGTGTGAGACCTTAATAGTGTGGCTCGACGCTCTAATGCTTAATAACACACGTAAGT	6182
Db	11323	TATTAAATTATAGCACTTAATGATTTTGTCTTCGACGATGAAGTACGACTAGACACGAAAGT	113820
Oy	6183	GACTGACGGGGGAACTGA--CATAGCGGTAAATCTGATGTACTTCCGAGGAAGCGTGGT	6241
Db	11383	GACCGCTACGCCCCAAATGATCCGACGCAAAATCGATGTACTTCCGAGGAACGTAGTGT	114420
Oy	6242	GCATATATGCCACGCGCGCT	6261
Db	11443	GCATATATGCATCAGCGTGGT	11462

RESULT 1.1

US-09-367-764-8
Sequence 8, Application US/09367764
Patent No. 6583121
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: System for the In Vivo Delivery and
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Selzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6583121th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09367,764
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 11703 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-367-764-8

Dh	5866	AAAGAAATTCATGCCCGCGTGTCTGCAACACGTCTGGAAGAGGAACCACTGAATCTACGTATACC	5925
Oy	760	AATATATGCGCCTCTGAAGAAATATGAAGCAGGTATCAATACGAAATGTAGAAATATATGA	819
Dh	5926	AGATGATGSCCAACCGAAGCCAAACAAAGATGTATACAGTCTCGTAAATGTAGAAATATGAA	5985
Oy	820	AAGCAATTCACGCGGACGACACTATTTCTGGAATTGGGCACATATCTATTCATTCAGAACTGA	879
Dh	5986	AAGCCATPACCACTGACGACGACTACTGTGACGACATGACGTGTATACCTGTGCCACAA--	6042
Oy	880	ATCTGTTCGAGTGTTCACAGAGTCAATATCTCTGACCAATCTACTCTGTCAACGGTAAATTA	939
Dh	6043	ATCAGCCAGATATGTATATGATCACTATCCGAATCCATTTGTACTCCAGTATAGCTACCGG	6102
Oy	940	ACAGTTTTACATCTGCAGAGGTGCGGTAAACGTGCAACTTAAGTTATCCAAAGAAATT	999
Dh	6103	CGAATCTATCCGATCCACAAGTTGGCTGTAGCTGTCTGTAAACAATATCTGATGAGAACT	6162
Oy	1000	ACCTTACATGATGCCAGTTATTTGTATACAGATGAATACGATGCTATCTTGACATGGTGG	1059
Dh	6163	ATCCGACATGATGACTCTTATCAGATTTCACGAGTTCAGATGCTTACTTGTGGATATGGTAG	6222
Oy	1060	ACGGGATGCTGTGCTCTGATATACGACGCACTTTTGTCCGGGCTAAACCTGAGAGTACACC	1119
Dh	6223	ACGGGACATGCTGCTGCTGTGATCTGCAACCTTTCTGCCCGCTTACCTTATGAAATTTACC	6282
Oy	1120	CAAGAAGCATAGTACTATTTGCAAGCCAGATATAGATCAGCCGTCCATCGCTATACAGA	1179
Dh	6283	CGAABAAAACATGATGATATAGAGCCCCGAATATCCGACATGGGGTTCCATACGACATGACA	6342
Oy	1180	ATATCATTTACAAATATGTATTTGGCTGACGCTACTAAAGAAATTTGCATGTTAACCCAAATGC	1239
Dh	6343	ACAGGCTACAAAATGTCTACTATTCGCCCACTAAABAAAATTTGCATGATCAGCAGATATGC	6402
Oy	1240	GAGAAATTAACCTGTCTTATGATTTCCGGCGGCACTTTATATGTGATTTGTTCAAGAAATAGCAT	1299
Dh	6403	GTGAATCTGCCAACACTGACTACGCGACATTTCAATGTGCAATGTCTTTGCAAAATATGCAAT	6462
Oy	1300	GCAATGATGATGATCTGGGATACCTTTTGGCATTAACCTTATTCGGCTAACTACAGAGAACG	1359
Dh	6463	GTATATGAGATATTTGGAGAGAGTTGCTGGAAAGCCAAATTAAGATTAACCATGATGTTTG	6522
Oy	1360	TTACGCAATATGTGACAAAGCTGAAGGGCCGAAAGCAGACGATTTGTTTCCGATATCTC	1419
Dh	6523	TCACCGCATATGTAGCTGATGAAAGGCCCTTAAGCCGCGCACTATTTTGCAAAAGACGT	6582
Oy	1420	ATATCTTAAACCTTGTCCAGAGAGTATACCATGATGATCAATTCGTATGATGATCTTAAAGAG	1479
Dh	6583	ATATATTTGTCTCCATTTCCAAAGAGTGTCTATGATATATTCGTATGACATGAAAGAG	6642
Oy	1480	ATGTCAAAAGTTATCTCCGCGACGAAACATACAGAGAGCGGCTTATAGTGCAGTTATTC	1539
Dh	6643	ACGTGAATGTTATKACACGAGCAGCAACAAACACAGAGAAACCCGAAGTACAGATGATAC	6702
Oy	1540	AGGCTGCAGATCCCCCTTGTACACGCTTACCTTTGCCGGATTCATCGGAAATTAGTCCGTA	1599
Dh	6703	AAGCCGAGAAACCCCTGGCAGCTGCTTACTTATGACGGGATTTACACGGGAATTAAGTGCTA	6762
Oy	1600	GACTGAATTCGGGCTCTTCTGTCCAAATATCCATACTCTCTTCGACATGTACAGCGGAATTT	1659
Dh	6763	GGCTTACGGCGCTTCTCTTCCAAACATTTCAACAGCTTTTGTGACATGTCCGGCGAGGATTT	6822
Oy	1660	TTGATGTGATTTATGTGGAACATTTTCCACACGCGGCAACCGATTTGGAAACGGAACATCG	1719
Dh	6823	TTGATGTGATTTATGTGGAACATTTTCCACACGCGGCAACCGATTTGGAAACGGAACATCG	6882
Oy	1720	CGTGTGTTGATTAAGCGAAGACGACGCTATTCGCAATTTCCGCGTGTGATGATCTCTTAGG	1779
Dh	6883	CATCATTTGCACAAAGCCAAAGACGACGCTATTTGGGTTAAACCGGTCTGATGATCTTTGAGG	6942
Oy	1780	ACTTAGGTGTGCACCAACCGCTCTTAGATTTGATAGAGCGGGGCTTTGGCAATATTCAT	1839
Dh	6943	ACCTGGGTGTGATCAACCACTACTGCACTTGATGTGATGATGATGATGATGATGATGATGATGAT	7002

QY 1840 CTGTGACCTTACTTACAGAGAGTTAAATTTGTGCCATGATGAATCCGGTAGT 1899
DB 7003 CCACCACTTACCTACGCGGTACTGTTTAAATTTGGGGCGATGTGAATCCGAAATGT 7062
QY 1900 TCTTAAACGCTGTGTGTCAACAACACTAGTCAATATCATGATTTGTAGCAGAGTACTACGTG 1959
DB 7063 TCTTCAACACTTTTGTCAACAAGTTTGAATGTGTTATGTCGACGAGAGTACTAGAAAG 7122
QY 1960 AACGTTAACACCGTCAAGCTGCGGGCTCTATGTGGGAGAGATAACATAGTGCATGTG 2019
DB 7123 AGCGCTTAAACGTCAGATGTGACGGCTTCATTGGGAGCAGCAACATCATACATGAG 7182
QY 2020 TCGTCTCCGACACCTTGTAGCGGAGAGATGCGCCACTTGGCTGAACATGGAAGTAA 2079
DB 7183 TAGTATCTGACAAAGAAATGGCTGAGAGGTGCGCACTGGCTCAACATGAGGTAAAG 7242
QY 2080 TTATTTAGTACAGTTATTTGTATCAAAAGCACTTACTTGTGGGGATTTATCTGTGTG 2139
DB 7243 TCATCGACGAGTACGCTGTAGAGAACCACTTACTTCTGCGGGGATTTATCTTGCAG 7302
QY 2140 ACCAGATAACAGGACACCTCTGCAGAGTCCGAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
DB 7303 ATTGCTTACTTCCACAGCGTGCCTGCGGAGCCCTGMAAGGCTGTTTAAGTTGG 7362
QY 2200 GAAAACCATTCGCTGATGATTAACCAAGACTGCGACCGCGCGGACATGCATGATG 2259
DB 7363 GTAAACCGCTCCACCGCAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7422
QY 2260 AAGCAATCGATGAGAAAGAAATTTGAAATTAAGAGATGAGAGAGAGAGAGAGAG 2319
DB 7423 AAACAGAGCGGTGGTTAGAGATGATTAACAGGACATTTAGCAGTGGCCGTGACAG 7482
QY 2320 GATACGAGATCACTGCGAGCGCTGATCATGCTCTGTCCAAGTTAGCCGAAAGCG 2379
DB 7483 GGTATGAGTGAACAATTAACAACCTGTCTTCTGATGAGCACTTTTCCGACAGCA 7542
QY 2380 TTAAGAACTTCAAGAGCATTAAGAGAGAGCCCAATCACCTCTGACGAGCTTAAATAG 2439
DB 7543 AAAGAGCTTCCAGAGCATCAAGAGGAGAAATTAAGCATCTCTGAGTGTCTTAAATAG 7602
QY 2440 TGACGTATGACACGCACTTAC-----CCACCGGAGAAATGTTTCCATACC 2486
DB 7603 CAGATAGTACATTTCACTGACTTAATTAACAACACACACACATGAAATGAGATCT 7662
QY 2487 CTCAGCTGAATTTTCCACAGTTTACCTTAACAATCGAGCTTACCGAATCCAAAC 2546
DB 7663 TTAAATGCTGCGCGCGCCCTTCCGCGCCCACTGCAATGTGAGAGCGCGAGAA 7722
QY 2547 CTCTAGGCGCGCTGAGGCGCTTTGGCCCCCGCTGCTGCTCAATCGAAGATCTTA 2606
DB 7723 GGAGGACGCGCGCCCGGATGCTGCTGCGCAACGCGCTGCTTCTTAAATTCAGCAGCA 7782
QY 2607 GGAGGTGATGATCACTTGAATTTCAACAACAGTACCTAATCCGCGCAGGTGCA- 2665
DB 7783 CCAAGCGCTGAGTGGCTTGAATTTGGAAGGCACTAAGACCTAACCCCACTCCAC 7842
QY 2666 -----CGCCAAAGAGAGAGAGTGTCTTAAAGCAAACTTACTAGCTTAAAG 2720
DB 7843 GCCCGCCACCGCGCAGAGAGAGAGCGCCCAAGCAACACCGAAGCGAAGAAACAA 7902
QY 2721 AGAAGCGAGCAAGCAAGAGAGAGAAAC---GCAAGCTTAACCGAGGAGAAAGAGCACT 2777
DB 7903 AAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7962
QY 2778 TGTGTATGAGATGAGAGTCCGACAGACATTTCCGATCATG---CTGAACGCGCAAGTGA 2834
DB 7963 TGGACATTAAGTTGAGAGCGGACAGATTGTTGAGAGTCAAGAGAGAGAGAGAGAG 8022
QY 2835 ATGATATGCTGCTGCTTCCGAGAGAGAGCTGATGAACCACTCCAGCTTAAAGAGAAA 2894
DB 8023 TCGGCGACGCACTGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8082

QY 2895 TTGATTAATGCAATTAAGCGCGCGTGAATTTGAAGAGGCTAGCATGTACGACTTTGAGT 2954
DB 8083 TCGACACCCCTGTGTGCTATCAAGAGCTCAAAATTTTACCAAGTCTGACAGATACAGATGAGT 8142
QY 2955 ACAGCGAGGTTCCCGAGAGATGAATTCAGACAGCTGAGTACACGAGCAAAACCAAC 3014
DB 8143 TCGCAGAGTTGCCAGTCAACATGAGAGAGTGAAGGACATTCACCTACACAGTAAACACCCG 8202
QY 3015 CGGCTTCTAACAATGCGGACCCACCGCGAGTCCAGTATGAGAAATGGAGATTTTACCGTAC 3074
DB 8203 AAGATTTATTAATCTGACACAGAGAGGAGTGAAGTGAAGTGAATTTTACATCC 8262
QY 3075 CGAGAGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3134
DB 8263 CTCGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8322
QY 3135 TTGTGCTATTTGTTCTAGAGAGTCAAAATGAGGAGACCGGTATCCGCGCTTTCAGTGTCA 3194
DB 8323 TTGTGCGATAGTCTCGGTGAGCTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 8382
QY 3195 CTTGGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3251
DB 8383 CTTGGAATAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8442
QY 3252 -----CACTAGTAAACAGCGCTATGCGTCTTGAATGTCAAGTCCCATGCGCAAAAC 3305
DB 8443 CAGACACATGCTGACAGGCAATGTGTTGCTCGAGAAATGATGCTTCCATGCGACCGCC 8502
QY 3306 CACCGGTGTCTATTCATGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
DB 8503 CGCCCAATGCTATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8562
QY 3366 ACAATCCAAATTAAGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
DB 8563 ACCATGAGGCTTAAGATACCTGCTCAATGAGAGAGAGAGAGAGAGAGAGAGAG 8622
QY 3423 CCAAGAGAGAGATTAAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3482
DB 8623 GCAAAAGAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8682
QY 3483 ATTGAGAGACTCAACGCGGTGTTTACGCGCCCAATTAATTAAGAGAGAGAGAGAG 3542
DB 8683 ACTGCAACATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8742
QY 3543 CTGATGATGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 3602
DB 8743 CGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8802
QY 3603 CTGAGAGATGACCAAAATTCGTTAATGCTTTCGACCAAGAGAGAGAGAGAGAGAG 3662
DB 8803 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8862
QY 3663 ACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3722
DB 8863 GACACATGAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8922
QY 3723 GGTACTTCTGTTAGTCAATGCTTCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3782
DB 8923 GATACCTTCTCTGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8982
QY 3783 GAGCATGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3842
DB 8983 GCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9042
QY 3843 AGAGATCTGTTTCCACCGCTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3902
DB 9043 AAAAATATGATCTACCTCCGTTTCAAGGTAAGAGAGAGAGAGAGAGAGAGAGAGAG 9102
QY 3903 TGAAGAGAGAGTCTGCGGAGTCAATTAACATGACAGAGAGAGAGAGAGAGAGAGAGAG 3962
DB 9103 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9162
QY 3963 CTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4022


```
Db 9163 CTAACCTGGAAGATCATCAGGGAAAGTTTACGAAAGCCGCTCTGGGAAGAACTTA 9222
Qy 4023 CTAACGAATTAAGTGGGCACTACAGCAAGGTATCTGAGACCGGAAAGAAATGA 4082
Db 9223 CGATAGGTGCAAGTGGGCACTACAGCAAGGTATCTGAGACCGGAAAGAAATGA 9282
Qy 4083 ACAGCTCAGTAAAGCAAAACAGTGCATTTGCTCAAGAGGCAACCAAGAAATGGTCT 4142
Db 9283 CTGGTTGACCGGCATCAGAGAGTGGCTCGCTTAAAGAGACCAACAAAGGGTCT 9342
Qy 4143 TCAACTCGCGGATCTTATTAAGCAACAGCACTAGTGAAGTAAATTCACATTC 4202
Db 9343 TCAACTCAGCGGATCTTATTAAGCAACAGCACTAGTGAAGTAAATTCACATTC 9402
Qy 4203 CATTCGCTTGAACCGAGAGTCTGCGGCTTAAAGTCAAGCTTACAGTCAAG 4262
Db 9403 CTTCGAAGTTGATCCCAAGTACCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 9462
Qy 4263 AGTGGTTCAAGGATCAGCTTCACTGATGATGATGATGATGATGATGATGATGATG 4322
Db 9463 ATGGCTTTAAACATATAGCTTCAATTAAGTATAGACCACTTGAATGCTTCAACCA 9522
Qy 4323 GAAATTTGGGCTGCGAGCAGACGCAACAGCAAGTGAATTAAGGTTTACATCAAG 4382
Db 9523 GAGAGTAAAGGAGCAACCCGGAACCAACCACTGAATGATGCTGGAAGAGCGTCAAG 9582
Qy 4383 ATTTTCTGTTGGGCGGAAAGGCTGAGTACGATGATGATGATGATGATGATGATGATG 4442
Db 9583 ACTTCAACGCTGAGCCGAGATGCTGAGTACATATGAGGAAATCATAGAGCACTGAGG 9642
Qy 4443 TCTGGGCGCAGAGTCCGAGCAGACGCAACCACTGATGATGATGATGATGATGATGATG 4502
Db 9643 TCTATGCGCAAGAGTCAAGCAGCAGAGAGACCTTCAAGATGCTTCAAGAAATATTAAGC 9702
Qy 4503 ACTATATCATGCGATCCAGTCTTACACTGCTATGCTGCTGCTGCTGCTGCTGCTGCT 4562
Db 9703 ATTTACTACATCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9762
Qy 4563 TCTGTTAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4622
Db 9763 TGATGATTTGGGCTGTAACCGTGGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 9822
Qy 4623 CGCATATGCGCTTGGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCA 4682
Db 9823 CGCATATGCGCTTGGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCA 9882
Qy 4683 TTGGGCAACCAACGCTGAAACATTTGGAAGAACTTGAACCATCTGCTGCTTAAACAC 4742
Db 9883 TTAGGTGCGCAATGCTGAAACGCTTCAACGCAACCATGATTAATTGCTGCAAGTGC 9942
Qy 4743 AACGTTTCTGCTGCGCAAGTGTGCACTTCTCTGCGCAAGCTTGTATCTGCTGCTGCT 4802
Db 9943 AGCGTTTCTGCTGCGCAAGTGTGCACTTCTCTGCGCAAGCTTGTATCTGCTGCTGCT 10002
Qy 4803 GCTTTTATGCTGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4862
Db 10003 GCTGCTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 10062
Qy 4863 CTTTGAACATGCAACCTGCTGCAATGTTCTGCGGATCCGCTATTAAGGCTGTTGCTG 4922
Db 10063 CTTTGAACATGCAACCTGCTGCAATGTTCTGCGGATCCGCTATTAAGGCTGTTGCTG 10122
Qy 4923 AACGCGAGGTTAGCGCGCACTTAACTGAGATCAAGGTCTCTCATCGGAATTAACAC 4982
Db 10123 AAAAGGAGGAGTATGCGCGCTCAATTTGAGATCACTGATCTCTCGGAGGTTTTCG 10182
Qy 4983 CTTGAATCAAGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5042
Db 10183 CTTGAATCAAGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 10242
Qy 5043 TTAATGCTGCGGCTCTGAGTGAAGGATCTCAAGGCGGATTAACATGCGCG 5102
Db 10243 TTAATGCTGCGGCTCTGAGTGAAGGATCTCAAGGCGGATTAACATGCGCG 10302
Qy 5103 TTTTGGCGGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5162
Db 10303 TCTTGGAGGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 10362
Qy 5163 ACAACACTAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5222
Db 10363 ACAACACTAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10422
Qy 5223 TCGCACTAAAGTTCACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5282
Db 10423 AGCGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 10482
Qy 5283 CCAACCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5342
Db 10483 CTAACGTTTCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10542
Qy 5343 AGGTATAGCAGGCGGCTATCAAGCGCTTCTTCAACCTTGAACATTAAGTCTGCTA 5402
Db 10543 AAGTATAGCAGGCGGCTATCAAGCGCTTCTTCAACCTTGAACATTAAGTCTGCTA 10602
Qy 5403 GAAAGGCGCTTCTTATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5462
Db 10603 ATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10662
Qy 5463 TCGGCGATTTCAAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5522
Db 10663 TTGAGACATTTCAAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10722
Qy 5523 GGTGCTGAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5582
Db 10723 GGTGCTGAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10782
Qy 5583 ATGAATGCTGAAGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5642
Db 10783 ATGAATGCTGAAGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10842
Qy 5643 AAATGAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5702
Db 10843 AAATGAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10902
Qy 5703 TTGACATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5762
Db 10903 TTGACATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10962
Qy 5763 GCTGCAAGTGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5822
Db 10963 GCTGCAAGTGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11022
Qy 5823 ACAAGCTGACAGGAGGCACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5882
Db 11023 ACAAGCTGACAGGAGGCACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11082
Qy 5883 AGGAAGCAGACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5942
Db 11083 AGGAAGCAGACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11142
Qy 5943 GCCCAAGCAATTTATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6002
Db 11143 GCCCAAGCAATTTATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11202
Qy 6003 GTAAACAGCGGCGACACATTAATGGAACACATTAAGTGAACCAAGAAATTCAGG 6062
Db 11203 GTAAACAGCGGCGACACATTAATGGAACACATTAAGTGAACCAAGAAATTCAGG 11262
Qy 6063 CGGCAAGTTTCAAAACATCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6122
Db 11263 CGGCAAGTTTCAAAACATCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11322
Qy 6123 TCAATGTTAGGACTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6182
Db 11323 TCAATGTTAGGACTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11382
```


QY 6183 GACTGAGCGGACACTGA-CATACCGGTAAACTGATGTTCTCCGAGGAGCGTGGT 6241
DB 11383 GACCGCTACGCCCCCATGATGATCGACGAGAAACTGATGTTCTCCGAGGAGCGTGGT 11442
QY 6242 GCATATGCGACGCGCGCT 6261
DB 11443 GCATATGATCATCGCGTGGT 11462

RESULT 12

US-08-801-263A-4

Sequence 4, Application US/08801263A

Patent No. 5811407

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Simpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

NUMBER OF SEQUENCES: 12 Expression of Heterologous Genes in the Bone Marrow

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell Seitzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 5811407th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,263A

FILING DATE: 19-FEB-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 11717 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-801-263A-4

Query Match 36.8%; Score 2386.4; DB 1; Length 11717;

Best Local Similarity 64.5%; Pred. No. 0;

Matches 3745; Conservative 0; Mismatches 1991; Indels 74; Gaps 9;

QY 580 ACTGACGATGAGACGGGAGCGGTATTTTCTCATCGGAAACAGGCCAAGTCACTTC 639
DB 5761 ACTGCTACCGGGGTAGGTGGTACATATTTTCACGAGACAGGCCCTGGGCACTTGC 5820
QY 640 AACAGAAATCAGTACGTCAATGTAACAAAGAACTATATGATGCGGCGGTCAG 699
DB 5821 AAATGAGATCCGCTTTCAGATACGCTTACAGAACCGACCTTGAAGCGCATGTTCTGG 5880
QY 700 AGAAGATTAACGCCCCCGCGCTCGATCTCGAAAGAGAGAAAATGTTACAGAAATCGC 759
DB 5881 AAAGATCTACGCCCCCGGTGCTCGACAGTCSAAAGAGAAAGAGCTCAAACTCAGTACC 5940
QY 760 AATATGCGCTCTGAAGAAATAGAGAGATATCAATCAGAAAGTAAATATGA 819
DB 5941 AGATGATCCCAACGAGCCAAACAAAGAGGTACAGCTTGAAGAAATGAGAAATCAGA 6000

QY 820 AAGCAATTACAGCGAGGAGACTCAATTTCTGATTTGGGACATATCTATCATCAGAAATGA 879
DB 6001 AAGCCATTAACCACTAGAGAGAGCTTTTCAAGGAGCTACGATGTAATCTCTGCACAG --- 6057
QY 880 ATCTGTGAGTGTATACAGAGTCAATTTCTGTACCAATCTACTCTGTCAACGGTAATTA 939
DB 6058 ATCAAGCAGAAATGCTATTAAGATCACTTACCCGAAACATCGTATTCAGACAGTACCGG 6117
QY 940 ACAGGTTACATCTGACAGAGGTGGGTTAAAGTCGACCTTATTCAGAGAAAT 999
DB 6118 GCACTTACTCTGACCCAAAGTTTGTCTGATGCTTTTGAACAACTATCTGACAGAAAT 6177
QY 1000 ACCCTACAGTACCGCAGTATTTGATATACAGATGATAGATGATGATCTTACATGATG 1059
DB 6178 ACCGAGGTGAGCATCTTATCAGATCAGACGAGATGATGATGATGATGATGATG 6237
QY 1060 ACGGCGATCGTCTGTCTGATATACAGCACTTTTGTCCGCTAACTGAGAACCTACC 1119
DB 6238 ACGGACAGTGGCTTGCCTAGATCTGCAACTTTTGTCCGCTAACTGAGAACCTACC 6297
QY 1120 CAAGAAACATAGCTATTTGACAGCCAGATGATGATGATGATGATGATGATGATG 1179
DB 6298 GAAAAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6357
QY 1180 ATACATTAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
DB 6358 AAGCTTCCAAAGAGTGTCTATGCTGCGGATGATGATGATGATGATGATGATGATG 6417
QY 1240 GAGATTAACCTGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1299
DB 6418 GTGATTTCCCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 6477
QY 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
DB 6478 GTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6537
QY 1360 TTACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419
DB 6538 TTACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6597
QY 1420 ATATCTTAAACCGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1479
DB 6598 ATATCTTAAACCGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 6657
QY 1480 ATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1539
DB 6658 ACGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6717
QY 1540 AGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
DB 6718 AAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6777
QY 1600 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1659
DB 6778 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6837
QY 1660 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1719
DB 6838 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6897
QY 1720 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779
DB 6898 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6957
QY 1780 ACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1839
DB 6958 ACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7017
QY 1840 CTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1899
DB 7018 CCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7077
QY 1900 TCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1959

Db	7078	TCCTCAGCCTCTTGTGTCAACA	CAGTTCTGAATGTCTGTTATGCGACGACGAGATTTGGAGG	7137
Oy	1960	AACGGTTAACACGCTCAGCGCTGCGCGG	CCTCTATCGGCGACGATTAACATATGCGATGGT	2019
Db	7138	AGCGGCTTAAACGTCGCAAAATGTG	CAGATTATTCGGGAGACACATCATACGAGG	7197
Oy	2020	TGCTCTCCGACACTTGTATGCGGAG	ATGSCGCCACTTGGCTGAACATGGAATAAAA	2079
Db	7198	TAGTATCTGACAAAGAAATGCTGAG	AGTGTGCCACTGGCTCAACATGAGAGTTAAG	7257
Oy	2080	TTATTTGATGCGATTATTTGGTAT	CAAGACCTTACTTTGTGGGGGATTTATCTGGTGG	2139
Db	7358	TCATTGAGCGAGTATCGGCGAGAG	ACGCTTACTTTGGGGTGAATCATCTTGGCAAG	7317
Oy	2140	ACCAAGTATACAGGCA	CAGCTCTGAGATCGACAGCCCTTAATAAAGCTTTTAAAGCTTG	2199
Db	7318	ATTGGGTATCTCCACAGCGTGTGCG	GTGGCGGACCCCTTGAANAAGCTGTTTAAATGG	7377
Oy	2200	GAAAAACATTTGCCAGTGCATGAT	ATACCACTGCGACCGCGCGGCACTGCATGATG	2259
Db	7378	GTAACACGCTCCACAGCGACGAC	GACAGAGAAAGAGAGCGCTCTGCTATGATG	7437
Oy	2260	AAGCAATCCGATGGAACGAATTTG	GAATTTAGGACGAGTTAGTGAAGGCGGTAGATCCA	2319
Db	7438	AAACAAAGCGCTGTTTGAAGTAA	GTAAACAGACCTTTGACAGTGGCCGTGCAATTC	7497
Oy	2320	GATACGAGATCATCTACTGCGACG	CGCTGATCATCGTCTCTGTCACAGTTACCGAAAGCG	2379
Db	7498	GGTATGAGGTAGACAAACATCA	CACCTGTCTGTCTGTGATTTAGAACTTTTGGCCAGACA	7557
Oy	2380	TTAAGAACTTCAAGAGCATTAAG	AGGAGGCCCAATCACCTCTAGCGCTGACCTTAATAG	2439
Db	7558	AAAGAGCATTTCAAGCCATCAG	AGGGGAAATTAAGCATCTCTACGCTGTCTTAATAGT	7617
Oy	2440	TGAGCTAGTA-----GACAG	CGACCTAACCCGCGAGAAATGTTTCCATACC	2486
Db	7618	CAGATAGCACATTTTCATCTGAT	CTAATACCAACACACACACATGATAGAGAAATCT	7677
Oy	2487	CTCAGCTGAACTTTCCACAG	TTTTAACCTTAACAAATCCGATGGCTTACCGAATCCAAACC	2546
Db	7678	TTTAACATGCTGGCGCGCCGCT	TCTCCGCGCCCACTGGCATTTGGAGGCGCGAGAA	7737
Oy	2547	CTCCTAGGCGCGCGTGAAGCG	GTGTTTGGCCCCCGCTGGCTCAATCGAAATCTTA	2606
Db	7738	GAGGACGAGGGGCCCCGATGCT	GCGCGCCGCAATGGGCTGGCTTCCCAATCCAGCACTGA	7797
Oy	2607	GGAGGTGATAGTCAACTTGA	CTTTCAACA-----ACGATCACTTAATCGCCGCCAG	2660
Db	7798	CCACAGCGGTGAGTGCCCTAG	TCTCATTTGACAGGCACTAAGACCTCAAAACCCCAAGCCAC	7857
Oy	2661	GTCCACGCGCCAAAGAAAGAA	AGATGCTCTTAAGCCAAACCTAATCAGCTTAAATAAG	2720
Db	7858	GCCCCGCGCCCGCCAGAAAG	ACGAGCGCCCAACCAACCGAAGCCGAAGAAACCA	7917
Oy	2721	AGAAGCAGCAACCCAAAGAGA	CGAAAC--GCAAGCTTAACCAAGGAAACGAAACGTA	2777
Db	7918	AAACACAGGAAGAAAGAAAG	AAAGCACTGCAAAACCCAAACCCGGAAGAGACACGTA	7977
Oy	2778	TGTGTATGAAGTTGAGT	CGGACAGACATTTCCGATC--ATGCTGAACGGCCAAAGTA	2834
Db	7978	TGGCACTCAAGTTGGAGGCG	CCACAGACTGTTCGACGTCAAAAATGAGGACGGAATGTCA	8037
Oy	2835	ATGCAATNGCTCGCTGTCTG	AGGAGAGGCTGATGAACCACTCACTTGAAGAAAAA	2894
Db	8038	TCGGGCAAGCACTGGCCAT	GATGAAGGAAGTAATGAACCACTCACTTGAAGAACTA	8097
Oy	2895	TTGATTAATGACAATTAAG	CGGCGCTGAATAATGAAGAGGCTAGCATGACACTTGGAGT	2954
Db	8098	TTGACACCCCTGTGTATCA	AAAGCTCAAAATTTCAACCAAGTCTGACATACGACATGAGT	8157
Oy	2955	ACGGGACGTTCCCGAAG	CATGAATTCGACACGCTGACGTACACAGCGACAAACCAAC	3014

Db	8158	TTGCACAGCTTGGCCGGTCTACATGAGAGTGAAGCCCTTCACTTACACAGCCGAACACCTCG	8217
Qy	3015	CGGGCTTTCTAACACTGTGCACACCGCGCAGTCCAGATGAGAGTGAAGATTTACCGTAC	3074
Db	8218	AAGGGTTTTCACATGGCACCACAGAGCGGTGAGATATGATGAGAGTGAATTTACATCC	8277
Qy	3075	CGAGAGAGTGGGCGGGGAAAGCGTCAACGCGGAAGCCATCTTGGACACAGAGGCAAG	3134
Db	8278	CCCCGGAGTAGAGAGCGAGAGGAGACAGTGTCTCCGATTTATGATTAATCTCAGGCCGG	8337
Qy	3135	TTGTGGCTATTGTTCTTAGAGAGGTGCAAAATGAGGGCGCGGTACGGGCTTTCAGTGTCA	3194
Db	8338	TTGTGCGGATATGTCCTTCGAGGGGCGCTGATGAGGAAACAAGAACTGCTCTTTCGATGTCA	8397
Qy	3195	CTTGGAAACCAAGAAAGGGGTGACCATTAGGGATACCCCGAAGTTCGTAGACCGTGT---	3251
Db	8398	CTGTGAATACCAAGAGGAAGACATTAAGACAAACCCCGAAGAGACAGAAAGTGTGTG	8457
Qy	3252	-----CACTAGTTTACAGCGCTATGCGTGTCTTCCGATGTCACTTTCCCATGCAAAAC	3305
Db	8458	CAGCACCACTGTCTACCGGCATGTCTTCTTGGAAACCTGAGCTTCCCATGCAATCGCC	8517
Qy	3306	CACCCGTGTCTATTCATCTGACCGCGAGAAACGAACCTGACGTGCTCGAAGAGACGTG	3365
Db	8518	CGCCCATATCTACACCCCGGAACCATCAAGAGCTTTGACATCTCTGAAGGAACGTGA	8577
Qy	3366	ACAATCCAAATTACGACACGCTGCTGGAGAAAGTCTTGAAT---GTCCATCACGCCGCG	3422
Db	8578	ACCACGAGGCTTAGACACCTGTCTCAAGCCATATTTGGGTGGGAGTGTGTCCGCAAA	8637
Qy	3423	CCAAACGAGCATTTACCGATGACTTCAACTGACCACTGACCACTCTTACTCTGGGGTTCTGCCGT	3482
Db	8638	GCAAAAGAACGTCACACTGACGACTTTTACCTTGACGAGCCCGTATCTTGGGCACTGCTGT	8697
Qy	3483	ATTGCACACACTGACGCGGTGTTTACGCCAAATTAATTAAGAACGTGGGGCGAAT	3542
Db	8698	ACTGTCACTCACTGAAACGCTGTTTAAAGCCGATTTAAGTACGAGAGGTCTGGGATGAAG	8757
Qy	3543	CTGATGATGATGATTTAGATCTCAGAGTCTCGGCAAAATTCGGCTTCAATCAGCAGCGCA	3602
Db	8758	CGGACGACAAACCTATACGATACAGACTTCCGCCAGTTTGGATACCAACCAAGCGGAG	8817
Qy	3603	CTGCGGATGTACACCAATTCCTTACATGTCTTTGACACGACCAATGACATCAAGGAAG	3662
Db	8818	CAGCAACTCAATTAAGTACCGGTACATGTCCGTGAGAGGATCATACCGTCAAGGAAG	8877
Qy	3663	ACAGATAGGAGAAATATAGTATATAGACATCTGAGACCTCGCGTGTCTTGGCCCAAG	3722
Db	8878	GCACTTAAGATGACATCAAGATAGGACCTCAGAACCGTGAAGAGGCTTGAAGTCAAAAG	8937
Qy	3723	GGTACTCTCGTTAGCTCAATGTCTCCAGGTGACAGTATACCGTCAATCAAGACG	3782
Db	8938	GATACTTTCTCTCCGAGAGTGTCTTCCAGGGGACAGGCTTAACGGTTAGTAGGAGTA	8997
Qy	3783	GAGCATGTGAATTCATGACACCGTGAAGAAAAAGATCAGAGGAAGTTTGTCCGTAGAG	3842
Db	8998	GCAACTCAGCAAGTATGACAAATGGCCCGCAAGTAAACAAATATTCGTGGGACGGG	9057
Qy	3843	AGGAGTACTTGTCCCACTCCGTTCATGGAAGAGTGTAAAGTCCACTTTACATCACT	3902
Db	9058	AAAAATATGAACCTATCCCGTTCACGGTAAAGATTCCTTTCACAGTGAACGACCGTCT	9117
Qy	3903	TGAAGGAGAGCTGTGCGCGGGTACATTAACATGACAGGCGCAGGCCCAACAGCGTATAGT	3962
Db	9118	TGAAGGAACAAACCGCGCTTACATCATATGACACGGCCGGGACCGACCGCTTACGT	9177
Qy	3963	CTATCTGGAAGAAAGCGTCAAGCGGAAGTGAATTAACCACTTCTGGCAAGAACGTCA	4022
Db	9178	CTATCTGGAAGGAATCATAGGGAAGTCTAAGCGGAAGCAACATTCGGAAAGAACATTTA	9237
Qy	4023	CTTACGAATGAAGTGTGGGACTACACGACAGGTATCTGTAGACACGGAAAGAAAGATGA	4082
Db	9238	CGTACGAGTCAAGTGTGGGGAATTCACAAACCGGTACCGTTTACGACCCGTACCGGAATCA	9297

QY 4083 ACGGCTGCATTAAGCAAAACAGTGCATTTGCTTACAGAGCGAGCAAAAGAAATGGGCT 4142
DB CGGGTGCACCGCCATCAAGAGTGCTGCTCTTAAGAGCGAACAAAGAGTGGTCT 9357
QY 4143 TCAACTCCCGGANTTTATTAGGACACAGACACTCAGTGCAGTAATTGCAATTC 4202
DB TCAATTCGCCGACTTGATCAGACATGCGACCAACGCGCCAAAGAAATTCATTTAC 9417
QY 4203 CATTCGCTTGACACCGACAGTGTCCGGTTCCGTTAGCTACACGCTTACAGTACAGA 4262
DB CTTCAGCTGATCCGAGTACCTGCAATGCTCCGTTGCCCCACGCGCGAAGCTAGTAC 9477
QY 4263 AGTGTGTAAGAGGACACCTCCACCTGACTGCAATGCGACCAATTTGCTGACAGCA 4322
DB ACGGCTTTAAACATCAAGCTTCCATTTAGACACAGACCACTGACATTTGCTACACCA 9537
QY 4323 GAAATTTGGGCTGCGAGCAAGCGACAGCAAGAAATGATTAACAGGGTCTTACATCAGGA 4382
DB GGAGACTAGGGGCAATCCGGAACCACTAGTATGATCATCGAAAGACGGTTAGA 9597
QY 4383 ATTTTCTGTGGGGGAGAAAGGCTGAGTACGTATGGGGTAAACATGAACCAAGTCAAG 4442
DB ACTTCACGTCGACCGAAGATGCGTGAATACATATGGGGCAATCAGAACCGGTAAAGG 9657
QY 4443 TCTGGGCCCAGAGTCGCGACACGAGCGACCCACATGATGCGCGCATGAGATCATATCC 4502
DB TCTATGCCCAAGAGTCTGACCAAGAGACCTTCAGGATGCGCACACGAAATAGTACAC 9717
QY 4503 ACTATTTATCATCGGATCCAGTCTTACACTGTCTATGTGTGTGTGTGTGTGTGTGTGT 4562
DB ATTACTACATCGCATCTGTGTACACATCTTAGCCGTGACATCAGTGTGTGTGTGTGT 9777
QY 4563 TCCGTGTAGGACATGATCATACAGACCTTGATGCGCAAGAGAAAGAAAGACTGCTGCA 4622
DB TGAATATTTGGGCTTAACTGTGACATTAATGTGCTGTAAGCGCGCGGTGTGCTGCA 9837
QY 4623 CGCCATAGCGCTTGCACCGAAGCGAAGCGTACCCACAGCATTTAGCGTTTGTGCTGCA 4682
DB CGCCATAGCGCTTGCACCGAAGCGAAGCGTATTCACATTTGCTGCGCACTTTGTGCTG 9897
QY 4683 TTTCGCAACCAACGCTGAAACATTTGAGAAATTTGAACCATCTGTGTGTTTAAACAAC 4742
DB TTAGGTGCGCTAATGCTGAACATTCACCGAGACATGATGTTACTATGTGTGAACAGCC 9957
QY 4743 AACGTTTCTGTGGGACAGTGTGTGATTTCTGTGGACGCTTGTATTTCTGTGCTGCT 4802
DB AGCCATTTCTGTGGGTCCAGCTGTGTATACCTGTGCGCTGTCTCATGCTTAAATGCGCT 10017
QY 4803 GCTTTTCATGCTGACCTTTTATTTTATTTGTTGACGAGGCTGCTGCGGGAAGGTAGAAG 4862
DB GTTGCTCATGCTGCTGCTTTTATTTTATTTGTTGCGGAGCTTACCTGCGGAAGGTAGAAG 10077
QY 4863 CTTTTCGAACTGCGACCACTGTGCCAAATGTTTCCGGGANTCCGTTATTAAGCGTTGTG 4922
DB CCTACGAACATGCGACCACTGTGCCAAATGTTGCCAAGTACCCTTAAGGACTTGTG 10137
QY 4923 AACGCGAGGTTTACGCGCACTTAACTGTGAATACAGGTGTCTCATCGGAATTAACAC 4982
DB AAAAGGCGAGGTTACGCGCACTTAACTGTGAATTAAGTATCTCATGCTTCCGAGGTTTTC 10197
QY 4983 CTTCACTAACAGAGTACGTAAGTACCTGCAATTCACACAGTCAATTCCTTCAACACAGAG 5042
DB CTTTTCACCAACAGAGTACATCACTGCAATTAACACAGTGTGATCCCTTCCCTTAAG 10257
QY 5043 TTAATGCTGCGGTCCTTCAGTGCAGAGCATCTTCAAGGCGGATTAACATGCGCG 5102
DB TCAATGCTGCGGTCCTTGAATGTCAAGCCGCGCTCAAGCAACATTAACCTGCAAG 10317
QY 5103 TTTTGGGCGGTGTACCTTTTCAATGTGGGAGGCGCAATGCTTCTGTGACATGAGA 5162
DB TCTTTGAGGGGTGTACCTTTTCAATGTGGGAGGAGCAATGTTTTCGACAGTGTAGA 10377

QY 5163 ACACACAATGATGAGGCGTACGTCGATTTGCTCCAGACTGACATATAGTACAGAG 5222
DB ACAGCCAAATGATGAGGCGTACGTCGATTTGCTCCAGACTGACATATAGTACAGAG 10437
QY 5223 TCGACATAAAGTTTCAACAGCTGCTTGAAGTGTGCGCTGTGTATTAATTAAGGCAACA 5282
DB AGGCAATTAAGTGTGATCTGCGGAGTGAAGTGTGAGTACGTATGATGTATGCGGAACA 10497
QY 5283 CCACCGGCGACCTGATACGTTTGTGCAATGCGTACAGCCAGGTTTCCACGCGGACTGGA 5342
DB CTACCAATTTCTTGAATGTGTATGAGTGAACGAGTACACAGGAACGTTTAAAGACTGA 10557
QY 5343 AGTTCATAGCAGGCGCATATCAGCGCTTTTTCACCTTTGACCATTAAGTGTGCTATCA 5402
DB AAGTCATAGTGTGACCAATTTTCAAGATGCTTTTACACATTTGATATCAAGTGTGATTC 10617
QY 5403 GAAAGGGCTTGTTTTACATACGATTTCTCTGATATGAGTATGAAACCAAGAGCGT 5462
DB ATCCGCGCTGTGTATCAATATGATCTTCCGGAATACGAGCGATGAAACCAAGAGCGT 10677
QY 5463 TCGCGATATTCAGACATCTTCCGCTTGAATGCTACAGACATATAGTACCCGCTGATAC 5522
DB TTGAGACATTTCAAGCTACCTCTTGAATGCTACCAAGATCTATGCGACAGACATTA 10737
QY 5523 GAGTGTGAGGCTTCTGTCAAGAACATCCAGTCCCTTCAACACCCAGACATATCAGGAT 5582
DB GACTACTAAGCTTTCGCGCAAGACGTGATGCTTCCGTAACAGCGAGCGCATTTGGAT 10797
QY 5583 ATGAAATGTGAGAACAACTCAGAGCGACCCCTGCAAGAAACAGACCATTTGGATGTA 5642
DB TCGAGATGTGAGAAACAACTCAGGCGCGCCACGTGACAGAAACCGCCCTTTGGGATGCA 10857
QY 5643 AAATTTGAATGAGGACCTTGTGAGACGCTTAACTGTGCTTAAAGGACATCTCTATCTGA 5702
DB AGATTTGAGTCAATCCGCTTGTGAGCGGTGAGCTGTCTAAGGAAACATTTCCATCTCTA 10917
QY 5703 TTGACATCTCTGATGACGCTTTTGTGAGATCATGAAATCAACAAATTTTGAAGTTA 5762
DB TCGACATCTCGAAGGCTGCTTTATCAGAGATCAGATGACACACTGATCTCAACAGTCA 10977
QY 5763 GCTGCACAGTACAGACTGATTTATTTCTGAGACTTTGTGTGTTCTTAACTTACAGT 5822
DB AATGTATGTAGTACAGTACGACTTACTACGAGGACTTGGGCGGATGCTACCTGCAAGT 11037
QY 5823 ACAAAGCTGACAGAGGAGGAGATTTGTCCATGCTACCTCCACGACAGCTGTTTGA 5882
DB ATGTATCGACCGGAGAGCAATGCTGTATCAATTCGCAATTCGAGCACAGCAACCTTCC 11097
QY 5883 AGGAAGGACACACATGTAAGTGCAGCGGTAGAGATTAACATTTTGAACATGCAATGCA 5942
DB AAGAGTGCACAGTTCATGTCTGAGAAAGAGGAGTACAGTACATTTCCAGACCGCA 11157
QY 5943 GCCCAAGCAAAATTTTATGATTTGCTATGCGGCAAGAGTCCACTGCAATGCTGAT 6002
DB GCCCAAGGCGAATTTATTTATGATGCTGTGTGTAAGAAACAAATGCAATGCAATGCAAT 11217
QY 6003 GTAAACCAACCGCGGACCACTAATTTGAGAAACCAATTAAGTGAACCAATTTCCAG 6062
DB GCAAAACCAACCGCTGACATTCGTGAGCAACCGGCAACAAATTAAGCAAAATTTCCAG 11277
QY 6063 CGGCAATTTCAAAACATCTTGAACAGTGTGCTTGTGACCTGTTTGGGGGAGATCATCC 6122
DB CCGCATTTCAAAACATTTATGAGTGTGCTGTGTTGCTTTTGGCGGCGCTGCTGCTG 11337
QY 6123 TCATTTGTGAGGACTTATGATGTTGTGCTGACGCTTATGCTTATTAACACAGCTAGAT 6182
DB TATTAATTTATGAGCTTATGATTTTGTGCTGACAGATGCTGATGACACAGAAAT 11397
QY 6183 GACTGAGCGGCGGACATGCA-CATAGCGGTAAACCTGATGATCTTCCAGAGAGGCTGT 6241
DB GACCGCTACGCGCCCAATGACCGAGCACAGCAAACTCGATGTACTTCCAGAGAACGTATGT 11457
QY 6242 GCATTAATGCCAGCGCGCT-----TGACACT 6268

1900 TCTTAAAGCTGTTGTTCAACACAGTCAATATCATGATTTGCTAGCAGAGTACTACGTG 1959
7078 TCCTCAGCCTCTTTGTCAACACAGTTCGTAATGTGGTATGCTGCGACAGAGATTTGAAGG 7137
1960 AACGGTTAACCAACGTCAGCGTCCGGGCTCTATGCGGACGATTAACATAGTGCATGTG 2019
7138 AGCGGCTTAAACGTCGCAAAATGTGACGCTTTATGCGGACGACCAACATCATACGAG 7197
2020 TCGTCTCCGACACTTGTATGTCGCGAGAGATGCGCACTTGCGTGAACATGAAAGTAAAA 2079
7198 TAGTATCTGACAAAGAAATGCTGAGAGGTGTGCCACTGCTCAACATGAGGTTAAGA 7257
2080 TTATATGACGATTAATGTTGATCAAGACCCCTACTTCTGTGGGGGATTTATCTGTGG 2139
7258 TCATGAGCGAGTCATCGCGCGAGACCGCTTACTTCTGCGGTGATTCATCTTGCAAG 7317
2140 ACCAATTAACAGGACAGCGCTGCAAGTGTGCAAGCCCTCTTAAAGGCTTTTAAAGCTTG 2199
7318 ATTGGTTACCTCCACAGCGTGTGCGTGGCGAGCCCTTGAAAGGCTGTTTAAAGTTGG 7377
2200 GAAAACCATTTGCGCATGTCATGATACCAAGCTGCGACCGCGCGCGGCACTGCATGATG 2259
7378 GTAAACCGCTCCAGCCGACGACGAGCAAGAGCAAGAAAGACGCGCTGTGCTAGATG 7437
2260 AAGCAATGCGATGTAACAGAAATTTGAATTTACGAGATTAAGTGAAGCCGTAGATCCA 2319
7438 AAACAAAGGCGTGTGTTAGATGATTAACAGACACTTACAGTGGCGGTGGCACTC 7497
2320 GATAGGATATCTACTGCGAGCGCTGATCATCACTCTCTGTCCAGTTAGCCGAAAGCG 2379
7498 GGTAGAGGTAGACAAATCATCACCTGTCTGCTGCAATGAGAACTTTTGGCCGAGACA 7557
2380 TTAAGACCTTCAAGCATTAAGAGGAGCGCAATCACTCTGACGGCTGACTTAATAGG 2439
7558 AAAGGACATTTCAAGCCATCAGAGGGGAAATTAACATCTCTACGGTGTCTTAATAGT 7617
2440 TGACCTAGTA-----GACAGCACTTAACCAACCGCGAGAAATGTTTCCATACC 2486
7618 CAGCATACACATTTTCATCTGATCTAATACCAACACCAACCACTGATAGAGATTTCT 7677
2487 CTCAGCTGAATTTTCCACCAAGTTTACCTTCAAAATCCGATGCGTTACCGAGATCCAAAC 2546
7678 TTAACTATCTCGGCGCGCGCCCTTCCGCGCCCACTGCAATGTGAGAGCGCGCGAGAA 7737
2547 CTCCTAGGCGCGCTGAGAGCGCTTTCGCGCCCGCTGCTGCTCAATTCAGATCTTA 2606
7738 GGAGCAGGCGCGCCCGATGCTGCTGCCCGCAATGCGCTGCTTCCAAATTCAGCAACTGA 7797
2607 GGAGTGCATAGTCAACTTGAATTTCAACA-----ACGATCACCTAATCCGCGCCGAG 2660
7798 CCAACAGCGTCAAGTCCCTAATGATTTGACAGGCAACTAGACTCAAAACCCACGCCCAAC 7857
2661 GTCCACCGCCAAAGAAAGAAAGTGTCTCTAAGCCAAAACCTTACGCTTAAAGAA 2720
7858 GCCCGCCCGCCGCGCAGAAAGAGCGCGCCCAACCAACCAAGCGAAAGCAAAACCA 7917
2721 AGAAGCAGCAAGCCAAAGAGAGCAAAAC---GCAACCTTAAACAGGAGAAAGCAAGCTA 2777
7918 AAACACAGAGAAAGAAAGAAAGCACTGCAAAACCCGAAACCCGAAAGAGACACTGA 7977
2778 TGTGATAGAGTTGGAGTGTGACAGACATTTCCGATC---ATGCTGAAGCGCCCAAGTGA 2834
7978 TGGCACTCAAGTTGAGGCGCAGACAGACTGTTCGAGCTCAAAAATAGAGAGCGAGATGCA 8037
2835 ATGATATGCTGCTGTGCGAGAGAGCTGATGAACCACTCCACGTTGAAGAGAAAA 2894
8038 TCGGCGCAGCACTGCGCATGAAAGAAAGTATGAACCACTCCACGTTGAAGAGACTGA 8097
2895 TTGATTAATGAGCAATTAGCGCGCGGTGAATTGAAGAAAGGCTAGCATGACATTTGGAGT 2954
8098 TTGACCACTCGTGTGATCAAAAGCTCAAAATTCACAAAGTGTGACAGATTAAGCAATGAGT 8157
2955 ACGGCGAGCTTCCCGAGAAATGAATCAGACGCTGCAAGTACCAAGCAAAACCAAC 3014

8158 TCGCAGATGCGCGGTCAACATGAGAGTGAAGCGCTTCACTTACACGAGCAACCCCTG 8217
3015 CGGCTTTTCAACATCTGACACCAACCGCGCATCTCAGTATGAGAAATGGAGATTTACCGTAC 3074
8218 AAGGTTTTTAACATCGGACCAACGAGCGGTGCAAGTATAGTGAAGTGAATTTTACATCC 8277
3075 CGAGAGAGTGGCGCGGAAAGCGGACAGCGGAAGCCGATCCTGGAACAACAGAGCAGAG 3134
8278 CCGCGGAGTGAAGAGAGAGAGACAGTGTGCTCCATTTATGATTAATCTCAGGCCGG 8337
3135 TTGTGCTAATTTGTTCTAGAGGTGCAATGAGGCGACGCGTACGCGCTTTCAAGTGTCA 3194
8338 TTGTGCGATAGTCTCTCGAGAGGCGGTGATGAGGGAACAAGAACTGCTCTTCGCGTCA 8397
3195 CTGGAACCAAGAAAGGCTGACATTAAGGATACCCCGAAAGTTCTGAACCGTGT--- 3251
8398 CCGGAATAGCAAAAGGGAAGCAATCAAGACACCCCGAAAGGGAACAGAGTGTGTG 8457
3252 -----CACTAGTTACAGCGGTATGCGCTTTTGAATGTCACTGCTCCATGCGCAAAAC 3305
8458 CAGACCACTGTGTACGCGCATGTGCTTTGTTGAAACGTGACCTTCCATGCAATCCGC 8517
3306 CACCGTGTCTATTCACTGACGCGCAAGACCACTGACGTGCTCGAAGAGACGTG 3365
8518 CGCCACATGTCTACACCGCGAACCATTCAGAGCTCTTGACATCTTGAAGAGACGTGA 8577
3366 ACAATCCAAATTAAGACACGCTGTGGAACGTCTTGAAT---GTCTATCAACCGCGC 3422
8578 ACCAGAGCGCTACGACACCTGCTCAAGCGCATATGCGGTGCGGATGTCTCCGCAAA 8637
3423 CCAACGAAAGATTAACGATGATTCACATGACAGCGCTCACTCGGGGTTTGCCTG 3482
8638 GCAAAAGAGCTCATGACACTTATCTTGAACAGCCGTACTTGAGGCAATGCTGT 8697
3483 ATTGCAAGACTCAACCGCTGTGTTACGCGCAATTAATTAAGAAACGTGTGGAAGCAAT 3542
8698 ACTGTACACATATGAAACCGTGTGCTTACCGATTAAGATGAGAGAGTCTGGAATGAAG 8757
3543 CTGATGATGATCGATTAAGATTCAGAGTCTGCGCAATTTGGCTCAATACAGGAGCA 3602
8758 CGGACGACAAACATACGATACAGACTTCCGCCCAATTTGATACCAACAAAGCGGAG 8817
3603 CTGGGATATGACCAAAATTCGGTATCATGCTTTTCGACACAGACCATGACATTCAGAGAG 3662
8818 CAGCAAGCTCAATTAAGTACCGCTTACATGTGCTGACAGAGATTAACCTCAAGAGAG 8877
3663 ACAGATGAGAAATAGCTATTCAGACATCTGACACCTGCGCTGCTTGGCCCAAG 3722
8878 GCATATGAGATGACATCAAGATCAGCACTAGAGACCGTGTAGAAAGCTTAGCTACAAAG 8937
3723 GGTACTTCTGTAGTCAATGTCTCTCAGAGTGAACAGTGAACCGTCAATCAAGAGCG 3782
8938 GATATCTTCTCTGCGAAGTGTCTCCAGGGGACAGGTAAAGTATGATTAAGAGATA 8997
3783 GAGCATCTGAGAAATTCATGACCGGTGAGAAAGAAAGATAGAGAGAAATTTGTGCTAGAG 3842
8998 GCAACTGACAACTCATGACAAATGCGCGCAAGATTAACCAAAATTCGTGGAGCGG 9057
3843 AGGATCTTGTCCCAACCGCTCATGAAAGTGTGAATGTCACAGTTCAGTCACT 3902
9058 AAAATATAGACTATCTCCGTTACGCTTAAGAAATTTCTTGCACAGTGTACACGCTC 9117
3903 TGAAGAGAGCTGTGCGGAGTACATAACATGACAGCGCAGGCCACACCGCTGAATAGT 3962
9118 TGAAGAAACAAACCGCGGTACTACATGACAGCGCGGACCGGACCGCTATAGT 9177
3963 CCTATCTGAGGAGCGCTCAGGCGCAAGTGTACATTAACCACTTTTGGCAAGAACGCA 4022
9178 CTTATCTGAGGAATCATGAGGAAAGTTCAGGAAAGCCCACTCCGAAAGAACATTA 9237
4023 CTTACGATGTATGTGTGCGCACTACAGCAGGTATCTGTAGACCGCAAGAGATGA 4082

D 9238 CGTACGAGTCAAGTGGCGGATTACAAGACCGGTACGTTACGACCCGTACCGAAATCA 9297
Q 4083 ACGGCTCACTAAAGCAAAACAGTCAATGGCTTCAAGAGGACCAACGAAATGGGCT 4142
D 9298 CGGGCTCACCGGCATCAGAGTGGCGTGGCTTAAAGAGCAACGAAAGGGGTCT 9357
Q 4143 TCAACTGCGGATCTTATTAGGACACAGACCACTCAGTCAAGTAAATTGCACTTC 4202
D 9358 TCATATGCGCGGACTTATGACATGCGACACAGCGCCCAAGGAAATTTGATTTAC 9417
Q 4203 CATTGCGCTTGACACCGACAGTCTGCCGGTTCGGTTAGCTCAGACGCTTACAGTCA 4262
D 9418 CTTTCAAGCTGATCCCAAGTACCTGATGGTCCCTGTGGCCACGCGCCCAAGCTAGTAC 9477
Q 4263 AGTGGTTCAAGGATACCTTCACTGACCTGAATGGCAACCAATGGTGAACA 4322
D 9478 ACGGCTTAAACATACATAGCTTCAATTAGACACAGACCACTGACATTTGTCACCA 9537
Q 4323 GAAATTTGGGCTGCGAGACGCAACAGAGAAATGATTAAGGCTTACATCCAG 4382
D 9538 GGAGACTAAGGGGCAATCCGGAACCACTAGTAATGATATGGAAGCGTTAGAA 9597
Q 4383 ATTTTCTGAGGCGAAGAGGCTGAGTACGATGGGTTAACCATGAACACTCAG 4442
D 9598 ACTTCACTGTCGACGAGATGGCTGGAAATATATGGGGCAATCAGAACCGGTTAAG 9657
Q 4443 TCTGGGCGCAGAGTCCGACACAGCCCAATGATGGCCGATGATCATATCC 4502
D 9658 TCTATGCCAAGATCTGACACAGAGACCTCAGATGGCCACAGAAATATATACAG 9717
Q 4503 ACTATTTCACTGCGATCCAGTCTACACTGCTATTTGCTGTGAGTCCGCTTGTCTA 4562
D 9718 ATTTACTACATCGCATCTCTGTATACACATCTTAGCCGTGCACTCAGCTGTGGCG 9777
Q 4563 TCTGTAGGACATGATCATCAGCACTTGCATGCCAAGCAAGAAAGACTGCTGA 4622
D 9778 TGATGATTTGGCGTAACTGTTGACACATTAATGCTGTAAGCCGCGGTAGGCTGCA 9837
Q 4623 CGCCATAGCGCTTGACCCGACCAAGCAACGTAACCAAGACTTAAAGCGTTTGTGCTGA 4682
D 9838 CGCCATATGCCCTGCGCCCAATGCCGATTTCCAACTTGCCTGACCTTTTGTGCTG 9897
Q 4683 TTGGGCAACCAAGCTGAAACATTTGGGAAACCTTTGAAACATCTGTGTTTAAAC 4742
D 9898 TTAGTGGGCTAATGCTGAACATTTCAACGAGACCATAGTATCTATGTCGAACGCC 9957
Q 4743 AACGTTTCTGTGGACACAGTTGTGATCTCTGTGGACGCGTTGTAATCTGTCCGT 4802
D 9958 AGCCATTTCTGTGGGTCCAGCTGTGTATACCCCTGGCGCTGTGATGTTCTAATGCGCT 10017
Q 4803 GCTTTTCACTGATGCTTTTATTTTATTTGTTGACAGCGTCTGCTGGGAAAGTATACG 4862
D 10018 GTTGTCTATGCTGCTGCTTTTATTTTATTTGTTGCTGGGCTTACCTGCGAAGTATACG 10077
Q 4863 CTTTGAACATGACACCACTGTGCCAAATGTTCCGGGATCCGTAATTAAGCGTTGCTG 4922
D 10078 CTTAGCAATGACACCACTTTTCCAAATGTGCCAGATACGTTAAGGCACTTGTG 10137
Q 4923 AACGCGAGTTTACGCGCACTTAACTGAGATCAGCGTGTCTCATGGAATTAACAC 4982
D 10138 AAAGGCGAGGTTAGCGCCGCTCAATTTGAGATTAATGTGATGCTCTGAGAGGTTTGC 10197
Q 4983 CTTCACTAACAAGAGTACGTAACCTGCAATTTCCACACAGTCAATTTCTTACACAG 5042
D 10198 CTTTCAACCAACCAAGATACATCCTGCAAAATTCACACATGTGGTCCCTCCCTTAAAG 10257
Q 5043 TTAATGCTGCGGCTCTGAGTGAAGGATCCTCAAGAGCGGATTAACATGCGCG 5102
D 10258 TCAATGCTGCGGCTCTTGAATGTCAAGCCGCGCTCAGACGACATATCTGCAAG 10317
Q 5103 TTTTGGCGGTGTACCTTTCATGTGGGAGGCGCAATGCTTCTGTACAGTGA 5162
D 10318 TCTTTGAGGGGTGTACCTTTCATGTGGGAGGAGCAATGTTTGGACAGTGA 10377

Q 5163 ACAACAAGTGTAGGCGGTAGCTGAGTTCCGCTCAGACTGCACTATATGATCAGCAG 5222
D 10378 ACAGCCAGATGATGAGCGGTAGCTGAAATTTGACAGATTTGCGGACAGCACCGCG 10437
Q 5223 TCGCACTAAAGTTCAACAGCTGCTGTGAAGTCCGCTGTATGATATGACGCA 5282
D 10438 AGCGATTAAGTGTACATGCGCGATGAAGTAAAGTAACTATATGTGTACGGGA 10497
Q 5283 CACCGGCACTGATACGTTTGTCAATGGCTTCAAGCCAGGTTTCTCAGCGGACCTGA 5342
D 10498 CTACAGTTTCTTGAATGTGTACGTGAACGAGTCAACAGGAACGTTAAAGACTGA 10557
Q 5343 AGGTATAGAGGCGGATATACGCGCTTTTACCTTTGACCAATTAAGCTGATCA 5402
D 10558 AAGTCTAGTGAACCAATTTTACGATCGTTTACACATTTGATCAGAGTGTATCC 10617
Q 5403 GAAAGGGGCTTGTTCACACTACGACTTCCCTGATATGAGGCTATGAACAGAGCGT 5462
D 10618 ATCGCGGCTGTGTACAACTATGATCTCCGGAATACGAGCGATGAACAGAGCGT 10677
Q 5463 TCGCGATATTCAAGCATCTCTGCTGTATGCTACAGACATGATAGCCGATGATAC 5522
D 10678 TTGAGACATTTCAAGCTACCTCTTGAAGTAAAGTCTCATGCGCAGACAGACTTA 10737
Q 5523 GGTGCTGAAGCCTTGTGTCAAGAACATCCAGCGCTTACACCCAGCAGATTCAGGGT 5582
D 10738 GACTCTCAAGCTTTCCGCGCAAGAGCTGATGCTCCGTACAGCAGCGCATCTGGAT 10797
Q 5583 ATGAATGTGGAAGAACCACTCAGAGACCCCTGCAAGAAAGCAACATTTGATGTA 5642
D 10798 TCGAGATGTGGAAGAACCACTCAGAGCGCCCACTGCAAGAAACCGCCCTTTGCGG 10857
Q 5643 AATTTGAATGAGGCTCTGCGAGCGTCTTAACTGTGTTACGGGCAATCCCTATCTGA 5702
D 10858 AGATTTGAGCAATCCGCTTGAAGCGGTGAGCTCTATAGGGAATTTCCATCTCTA 10917
Q 5703 TTGACATCCCTGATGACGCTTTTGTGATCATCAGATCAACCAACATTTTGAAGTGA 5762
D 10918 TCGACATCCGAAACGCTGCTTTATACAGACATCAATGACCACTGTGTCAACAGTCA 10977
Q 5763 GCTGCAAGTACAGACATGATTAATTTCTGACAGCTTTGTGTTCTTAACTACAGT 5822
D 10978 AATGTATGTCAAGTATGATCTTACAGCGACTTCCGCGGATGATACCTGAGT 11037
Q 5823 ACAAGCTGACAGGAGGACATTTGCTCACTTCCACTCCAGACAGCTGTTTGA 5882
D 11038 ATGTATCCAGCGCGAAGGACCAATGCCCTGTATTCGATTCGAGACAGCAACCTCC 11097
Q 5883 AGGAAGGACCAACATGATGACGCGGTAGGACAGATTAACATTAATTTAGCATGCA 5942
D 11098 AAGATGCAAGTTCACTGCTGGAAGAAAGAGCGGTGACAGTACCTTACAGCCGCA 11157
Q 5943 GCCCACAAGCAATTTTATATGTTTCCGTATGCGGCAAGAACTCACTGCAATGCTGAAT 6002
D 11158 GCCCACAAGGAACTTATATGTTATGCTGTGTGTGAAGAAACAATGCAATGCAAGAT 11217
Q 6003 GTAAACCAACGCGGACCAATTAATGAGAACCACTAAAGTGCACCAAGAAATTCAGG 6062
D 11218 GCAAAACCAACGATGACATATGATGAGACCCCGCAAAATGACCAAGAAATTCAG 11277
Q 6063 CGGCACTTCAAAAACATCTTGAACCTGCTGTGACGTTTGGGGAGCATCATCC 6122
D 11278 CGGCACTCAAAAACATCTTGAAGTGTGCTGCTGCTTTCGCGGCGCTCTGCTGC 11337
Q 6123 TCATTTGTAGACATTAATGATGTTGTCTGACGCTTATGCTTATTAACACAGTAT 6182
D 11338 TATTAATTAAGAGCTTATATGTTTGTGACAGATGATGATGATGACAGAGAT 11397
Q 6183 GACTGACGCGGACACTGA-CATAGCGTAAACTGATGATCTTCCGAGAAAGCGTGT 6241
D 11398 GACCGTACGCGCCCAATGACCCGACAGCAAACTCGATGATCTTCCGAGAACTGATGT 11457

Qy 6242 GCATATCCACGCCCT-----TGACACT 6268
Db 11458 GCATATCCACGCCCTGATATATAGATCCCGCTACCGCGGCAATATAGCAACCC 11517
Qy 6269 AAAATTCATGATTTCCGAGGAGACAGTCATTAATGCTGTGACGT 6318
Db 11518 AAAATTCATGATTTCCGAGGAGACAGTCATTAATGCTGTGACGT 11567

RESULT 14
US-09-367-764-4
Sequence 4, Application US/09367764
Patent No. 6583121
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6583121th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/367,764
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11717 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-367-764-4

Query Match 36.8%; Score 2386.4; DB 4; Length 11717;
Best Local Similarity 64.5%; Pred. No. 0; Mismatches 1991; Indels 74; Gaps 9;
Matches 3745; Conservative 0;

Qy 580 ACTGACGATGAAAGCGGATATTTCTCATCGGAAACAGGCCAGGCTACCTTC 639
Db 5761 ACTGACGATGAAAGCGGATGATGATATTTCTCATCGGAAACAGGCCAGGCTACCTTC 5820
Qy 640 AACGAAATCACTAGCTCAATGTAACTAACAAGAACCTTAATTTGATGAGGCGCTGCATG 699
Db 5821 AAATGAGATCGCTTGTGCAATCACTTACGAAACGACCTTGAAGGCGCAATGTTCTGG 5880
Qy 700 AGAAGTATTAAGCGCGCGCTGATCTGAAAGAGAAATGTATTAAGAAGAACTGC 759
Db 5881 AAAGATTAAGCGCGCGCTGCTGCAAGTGAAGAGAAAGAGCTCAAACTCAGATACC 5940
Qy 760 AATATGCGCTCTGAAAGAAATAGACAGGATATCAATCAAGAAAGTAGAAATATGA 819

Db 5941 AGATGATCCCAACCGAACCAACAAAGCAGTACCACTATAGAAAGTAGAAATATCA 6000
Qy 820 AAGCAATTAACGAGGAGCACTATTTGATTTGGGACATATCATCATGAAAGTA 879
Db 6001 AAGCAATTAACGAGGAGCACTATTTGATTTGGGACATATCATCATGAAAGTA 6057
Qy 880 ATCTGTGAGTGTATTAAGATCAATTAATCTGTACCAATCTACTGTCAACGATATTA 939
Db 6058 ATCAAGCAAGATGTATTAAGATCAATTAATCTGTACCAATCTACTGTCAACGATATTA 6117
Qy 940 ACAGTTTACATCTGCAAGAGTGGCGGTTAAACGTCACCTAATTAATCAAGAAAT 999
Db 6118 CGAATCTATGACCCAAAGTTTGGTGTAGCTTTGCAACAACTATCTGATGAGAAAT 6177
Qy 1000 ACCCTACAGTACCAAGTATTTGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1059
Db 6178 ACCGACGATGAGTATTTATCAAGTACAGGAGTATGATTTGATTTGATTTGATTTG 6237
Qy 1060 ACAGGATCTGCTGTGATGATTAAGTACAGGAGTATTTGATTTGATTTGATTTGATTT 1119
Db 6238 ACAGGATCTGCTGTGATGATTAAGTACAGGAGTATTTGATTTGATTTGATTTGATTT 6297
Qy 1120 CAAGAACATAGCTATTTGACAGGAGTATTAAGTACAGGAGTATTTGATTTGATTTG 1179
Db 6298 CAAGAACATAGCTATTTGACAGGAGTATTAAGTACAGGAGTATTTGATTTGATTTG 6357
Qy 1180 ATACATTAAGATTAAGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1239
Db 6358 ATACATTAAGATTAAGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 6417
Qy 1240 GAGATTAAGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1299
Db 6418 GAGATTAAGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 6477
Qy 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
Db 6478 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6537
Qy 1360 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419
Db 6538 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6597
Qy 1420 ATTAATCTAAGACGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1479
Db 6598 ATTAATCTAAGACGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 6657
Qy 1480 ATGTCAAAGTATCTCCGAGCAAGAAACATTAAGAGAGAGAGAGAGAGAGAGAGAG 1539
Db 6658 ATGTCAAAGTATCTCCGAGCAAGAAACATTAAGAGAGAGAGAGAGAGAGAGAGAG 6717
Qy 1540 AGGCTGAGATCCCTTGTCTACCGCTTACCTTTGCGGATTCATTCGGAATTAATGTC 1599
Db 6718 AGGCTGAGATCCCTTGTCTACCGCTTACCTTTGCGGATTCATTCGGAATTAATGTC 6777
Qy 1600 GACTGAATGCGGATGCTTGTGCAATATCAATCAATCAATCAATCAATCAATCAATCAAT 1659
Db 6778 GACTGAATGCGGATGCTTGTGCAATATCAATCAATCAATCAATCAATCAATCAATCAAT 6837
Qy 6838 TTGATGCAATATGAGCAAGAACATTTCAAGAGGAGAGAGAGAGAGAGAGAGAGAG 6897
Db 1720 TTGATGCAATATGAGCAAGAACATTTCAAGAGGAGAGAGAGAGAGAGAGAGAGAG 1779
Qy 6898 CTTGCTTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 6957
Db 1780 ACTTGAAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1839
Qy 6958 ACTTGAAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 7017
Db 1840 CTGTGACATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1899
Qy 7018 CCACCAATCTGCGGAGTACCGGTTCAATTCGAGGAGATGATGAAATCCGGAATGT 7077

Qy	1900	TCCTTAAGCGCTGTTGTCTGAACACACTAGTCAATATCAATGATGTGCTAGACGAGTACTACTG	1953
Db	7078	TCCTCAGCGCTCTTGTGTCAACACAGATTTCTGAATGTCTGTAATGTCTGCAACGAGATATTGGAG	7137
Qy	1960	AACCGTTAAACACAGCTCAGCGCTGCGCGGCTCTATCTGGCGACGATACATAGTGCATG	2019
Db	7138	AGCGCGTTAAACGCTCCAAATGTGTCAAGATTTATTCGGGACGACACATCAACGAG	7197
Qy	2020	TCGCTCTCCGACCTTGATGGCGGAGAGATGCGCACTTGGCTGAACATGGAAATAAA	2079
Db	7198	TAGATATCTGCAGAAAGAAATGGCTAGAGGTGTGCACCTGGCTCAACATGAGGTTAAGA	7257
Qy	2080	TTATTTGATGCGAGTAAATGATGATCAAAAGCACCTTCTCTGAGGGAGATTATCTCGTGG	2139
Db	7258	TCATTTGACGCGAGTCAATGGGCGAGAGACCGCTTACTTCTGGGTGGATTAATCTTGCAG	7317
Qy	2140	ACCAGATPACAGGCGACAGCCCTGACAGAGTCCGACAGCCCTTAATAAAGCTTTTAAGCTTG	2199
Db	7318	ATTGCGATTACTCTCACAGCGCTGTGCGGTGGCGGACCCCTTGAAAAGGCTGTTTAAGTTGG	7377
Qy	2200	GAAACCAATTTGCCAGTGGATGATACCCAGACCTGGACCCGCGCGGCAATGCAATGATG	2259
Db	7378	GTAACACCGCTCCAGCGGACGACGACGAGACGAAAGACAGGCGCTCTGCTTAATG	7437
Qy	2260	AAGCAATGCGATGAGAACAGAAATTTGAAATTACGACGAGTTAGTAAAGCCGTAAATATCA	2319
Db	7438	AAACAAAGCGCGTGTTTAGATAGATATGATATACAGACACTTACAGTGGCGGTGGCACTC	7497
Qy	2320	GATACGAGATCATCTGGCAGGCGCTGATCATCAGTCTCTGTCCAAGTTAGCCGAAACG	2379
Db	7498	GGTATGAGGTAGACAAACATCAACACTGTCTCTGGCAGTTAGAACTTTTGCCAGAGCA	7557
Qy	2380	TTAAGAACTTCAAGAGCCTAAGAGGAGGCCAATCAACCTCTAGCGCTGACCTTAATAG	2439
Db	7558	AAAGAGCATTTCAAGCCATCAAGGGGAAATAAAGCATCTTTAGGTGTCTTAATAGT	7617
Qy	2440	TGACGTAGTA-----GACACGACCTAACCCACCGGAGATGTTTCCATACC	2486
Db	7618	CAGCATAGCATTTTCATCTGACTAATATCAACACACACACATGAAATAGAGATTTCT	7677
Qy	2487	CTCAGCTGAACTTTCCACCACTTTAACCTTAACCTTAACATTCGATGGCTTAACGAAATCAAC	2546
Db	7678	TTAAACATGCTGGGCGCGCCCTTCCCGGCCCCCATCTGCATATGGAGGCGCGGAGAA	7737
Qy	2547	CTCCTAGGCGCGCGTGAAGCGCGTTTGGCCCCCGCTGGCTGCTCAATCGAAGATCTTA	2606
Db	7738	GGAAGCAGGGCGGCCCGATGCTCGCGCAATGGGTGTGCTTCCAAATTCAGCAACTGA	7797
Qy	2607	GGAAGTGCATGTCACTTGACTTTCAAAACA-----ACGATCACTAATCCGCCCGCAG	2660
Db	7798	CCACAGCGCTGAGTGCCTTAATTTGACACGCGCACTGACCTTCMAACCCCAAGCCAC	7857
Qy	2661	GTCCACCGCCAAAGAAAGAAAGATGTCTCTTAAGCCCAAAACCTACTCAGCCTTAAAGAA	2720
Db	7858	GCCCCGCGCGCCGCGCAAGAAAGACAGCGGCCCAAGCAACACCGAAGCCGAAGAAACCA	7917
Qy	2721	AGAACGACGAACCAAGAGACGAAC--GCAGCCTTAAACAGAGAAACGACAACTGA	2777
Db	7918	AAACACAGGAGAAAGAAAGAAAGCACTGCAAAAACCAAAACCGGAAAGAGACAACTGA	7977
Qy	2778	TGTGTATGAATTTGAGATCGGACAAAGACATTTCCGATC--ATGCTGAACGGCCAAAGTA	2834
Db	7978	TGGCACTTAAATTTGAGGCGCACAGACTGTTGACGTAAAAATGAGGACGAGATGTGCA	8037
Qy	2835	ATGGAATATGCTGCTGTCCGAGGAAAGCTGATGAACCACTCCACCTTGAAGAAATAAA	2894
Db	8038	TCCGCGACGCACTGCCCATGGAAGGAAAGATATGAACCACTCCACCTGAAGAAAGACTA	8097
Qy	2895	TTGATATATGACAAATTAGCGGCGGTGAATTTGAAGAGGCTTAGCATGTTCGAGT	2954
Db	8098	TTGACCACTCTGTGTATCAAGGTCAAAATTCACCAAGTGTGTGACATATGACATGTGAGT	8157

QY	2955	ACGGCGACGTTTCCCCAGAAATGAAATTCACACACGCTGCAGTTCACACGAGGACAAACAC	3014
Db	8158	TGCGACAGTTGGCCGGTCAACATGAAAGTAGAGGGTTTCACTTACACAGCGAACACCTTG	8217
QY	3015	CGGGCTTCTCAACCTGGCAACAACGGCCAGTCCAGTATGAAATGGAGATTTTACCGGTAC	3074
Db	8218	AAGGTTTTCACATGGCAACACGAGACCGGTGCGATATGATGGAGATGAAATTTACATCC	8277
QY	3075	CGAGAGAGTGGGGGGAAGGCGACAGCCGAAAGCCGATCTTGGACAACAGAGGCAAG	3134
Db	8278	CCCGGGAGTGGAGGCAAGAGAGACAGTGGTCTCCGATTTATGATTAACCTACAGCCGGG	8337
QY	3135	TTGTGGCTATTGTTCTTAGAGGTGCAATATAGGGCAGCGGTACCGGCCCTTCAAGTGTCA	3194
Db	8338	TTGTTCGGATATGTCCTCGAGGGGCGTAGTAGGGAAACAAGACCTGCTTTCGGTCTGTCA	8397
QY	3195	CTTGGAAACAGAAAGGGGTGAACAATTAGGGATACCCCGAAGGTTCTGAACCCGGTGT---	3251
Db	8398	CTTGGAAATGCAAAAGGGAAGACATCAAGACACCCCGAAGGGAACAGAAAGTGTCTG	8457
QY	3252	-----CACTAGTTACAGCGGTATGCGTCTTTGAAATGTCAAGTTCCCATGCAAAAC	3305
Db	8458	CAGACCAACCTGTCAACGGCCATGTGCTTGTGGAAAGTGAAGCTTCCCATGCAATGACC	8517
QY	3306	CACCCGTGTCTATTCTACTGACGCCAGAACGAACACTCGACGTCTCGAAGAGAACGTGG	3365
Db	8518	CGCCACATGTCTAACCCGCGAACATCCAGAGTCTTGTGACATCTTGAAGAGAACGTGA	8577
QY	3366	ACAATCCAAATTAAGACACACGCTGTGGAGAACGTCTTGAAT---GTCCATACACCGCGC	3422
Db	8578	ACACAGAGCCTTAGACACCTTGTCTCAAGCCATATTTGGGTGGGATCGTCCGGCAGAA	8637
QY	3423	CCAAACGAAGCATTACCGATGACTTCAACCTGACAGTCCCTACCTGGGGTTCTGCCGT	3482
Db	8638	GCAAAAGAGGCTCACTGACGACCTTTACTTGAACAGCCGCTTCTTGGGCAATGCTCGT	8697
QY	3483	ATTGCAGACACTCAACCGCTGTCTTCAAGCCCAATAAAAATTGAGAACGTGTGGACCAAT	3542
Db	8698	ACTGTCAACCAATCACTGAACCGTGTCTTGTAGCCGATTAAGATCGACAGAGTCTGGATAGAG	8757
QY	3543	CTGATGATGATGATGATTAAGATTCAGAGGTCTGGGACAAATCGGCTTACATCAGGACAGCA	3602
Db	8758	CGGACGACACACCAATAGCATACAGACTTTCGCCCACTTTGGATAGACCAAGCGGAG	8817
QY	3603	CTGGCGATGTCAACCAAAATTCGTTACATGTCTTTGACACAGACCATGACATCAAGAGAG	3662
Db	8818	CAGCAAGCTCAAAATTAAGTACCGCTACATGTGCTCGAGCAGGATCATACCGTCAAGAGAG	8877
QY	3663	ACAGTATGGAAAAATAGCTATACAGCACAATTTGGAACCTCGCCGTGCTTTGGCAAAAG	3722
Db	8878	GCACATATGATGACATCAAGATTCAGACCTCAGGACCGGTGTAGAAAGCTTAGCTACAAAG	8937
QY	3723	GGTACTTCCGTAGCTCAATGTCTCCAGGTGACAGTGAACCGTCAGTATCAAGAGCG	3782
Db	8938	GATATCTTCTCCTCGCAAGTGTCTCCAGGGGACAGGTATACGTTATGATACGAGTA	8997
QY	3783	GAGCACTGGAATTTATGACACCGTGGAAAAAAGATCAGAGGAGATTTTGTCCGTAGAG	3842
Db	8998	GCAACTGACGACAGTCATGACACATTTGGCCGCAAGATTAACAAAAATTCGTGGAGAGGG	9057
QY	3843	AGAGTACTTGTCCACCCGTCATAGGAAGCTGTGAAGTGCACGTTTATGATCACT	3902
Db	9058	AAAAATATGACCTTACTCTCCGTTCACGGTATGAAGATTTCTTGCACAGTGTACACCGTCT	9117
QY	3903	TGAAGAGACGTCTGCGGGGTACATTAACAGACAGGCCAGGGCCACAGCCGATTAAGT	3962
Db	9118	TGAAGAAACAACCGCGGTACATCACTATGACAAGCCGGGACCGGACGCTTATAGT	9177
QY	3963	CTTATCTGAGGAAGCGTCAGGCAAGTGTACATTTAAACAACCTTCTTGGCAAGAACGTCA	4022
Db	9178	CTTATCTGAGGAATCATCAGGAAAGTCTACGGGAACCAACATCCGGAAGAAACATTA	9237
QY	4023	CCTTAAATGTAAATGTGGCGACTTACAGCACAGGATATCTGTAGACACCGGAACGAAGATGA	4082

Db	9238	CGTACGATGCAAGTCCGGCGATTACAGACCGGTAACGGTACCGATCCGTAACCGAATCA	9297
Oy	4083	ACGGCTGCATTAAGCAAAACAGTGCATTTGCTCTCAAGAGGCAACAAACGAATGGGCT	4142
Db	9298	CGGGCTGCACCGCCATCAAGCAGTGGCTCGCTTAAGACGACCAACGAAAGTGGCT	9357
Oy	4143	TCAACTCGCCGGATCTTTATTAAGCACACAGACCACTCAAGTGCMAAGTAATTCACATTC	4202
Db	9358	TCAAATTCGCCCGACTTGATCAGACATGCGCCACACACGGCCCAAGGGAAATTGCAATTAC	9417
Oy	4203	CATTCGGCTTGAACCGACAGTTCGCCGGTTCCGTTAGCTTACACGGCTTACAGTCACGA	4262
Db	9418	CTTTAAGCTGATCCCGAGTACTGTGCATAGTTCCTGTGGCCACGCGCGCAACGTATAC	9477
Oy	4263	AGTGGTTCAAAAGGCATCAACCTCTCACTGACTCAATGCAACGCAACATTTGCTCAACGA	4322
Db	9478	ACGGCTTTAAACATCAAGCTCTCCAAATTAGACACAGACCACTGACATTTGCTCACACCA	9537
Oy	4333	GAAAAATGGGGCTCGACACAGCAACGACAGAAATGATTAACAGGCTCTACATCCAGA	4382
Db	9538	GGAGACTAGGGGCAAAATCCGGAACCAATCTGAATGATCATGGAAGAAGCGTTAGAA	9597
Oy	4383	ATTTTCTGTGGGGGGAAGGGCTGAGAGTACGTATGGGGTAAACATTAACCACTCAGAG	4442
Db	9598	ACTTACCGTGCACCGAATGCGCTCGAATAATATGGGGAAATCAACACCGGTAAAGG	9657
Oy	4443	TCGCGGCCAGGAGTCCGACCAAGCGCAACCAATGATATGGCGCGCATGATATCATATCC	4502
Db	9658	TCTAAGCCCAAGAGCTGACCAAGAGACCTCTACAGGATGGCCACAGAAATAGTACAGC	9717
Oy	4503	ACTATTATCATCGGCATCCAGTCACTGATATTTGTCGTGTGTCGTGTCGTCTTGCTA	4562
Db	9718	ATTACTACCATCGCATCTGTGTGACCAATCTTAGCGTGCATCAGCTCTGTGGGA	9777
Oy	4563	TCCGTGTAGGCACATGATCAGACGTTGATGCGCAAGCAAGAAAGACATGCTGCA	4622
Db	9778	TGATGATTTGGGTTAACTGTTCGAGCATTAATGTGCTGTAAAGCGCGCTGATGCGCTGA	9837
Oy	4623	CGCCATAAGCGCTTGCACCGAAGCAACGATACCAACAGCATATGCGTTTGTGCTGCA	4682
Db	9838	CGCCATAAGCGCTTGCACCGAAGCAACGATACCAACAGCATATGCGTTTGTGCTGCA	9897
Oy	4683	TTGCGCCAAACGCTGAAACATTTGGAGAACTTTGAAACATCTGTGTTTAAACAC	4742
Db	9898	TTAGGTCCGCTAAATGCTGAAACATTTCAACGAGACATGATGTTAATGATGCGAAGCC	9957
Oy	4743	AACGTTTCTCTGGGCAAGTTTGTGCAATCTCTGTGCAAGCGCTTGTATTTCTGTTCCGT	4802
Db	9958	AGCCATTTCTTGGGTTCAGCTGTGTATACCCCTGGCGCTGTATCATCTTTAATGCGCT	1001
Oy	4803	GCTTTTCAAGCGCATGCTTTTATTTATGTTGTGACGGGCTGCGCTGGGGAAGTAAAGC	4862
Db	10018	GTTGCTATGCTGCTGCTGCTTTTATTTATGTTGTGCGGGCGCTTACCTGGCAAGGTAAAGC	10077
Oy	4863	CCTTGAACATGCGACCACTGTGCCAAATGTTTCGGGGATCCCGTATTAAGCGTTGTGCG	4922
Db	10078	CCTAGGAACATGCGACCACTGTGCCAAATGTTTCGGGGATCCCGTATTAAGCGTTGTGCG	1013
Oy	4923	AACGGCGAGTTTACCGCCCATTTAACTGTGAGATCAGCGTGTCTCATCGGAATTAACAC	4982
Db	10138	AAAAGGCGAGGTACCGCCCGCTCAATTTGGAGATTACTGTATGTCTCCGAGGTTTGTG	1019
Oy	4983	CTTCAACATAACAGAGATGACGATCCTCAAAATTCACACAGTCAATTCCTTCAACACAG	5042
Db	10198	CTTCAACATAACAGAGATGATCACTTCAAAATTCACACAGTCAATTCCTTCAACACAG	1025
Oy	5043	TTAAATGCTGCGGCTCCTCGAGTGCAGAGGATCCTCAAAAGCGGATTTACATGCGCG	5102
Db	10258	TCAATATGCTGCGGCTCCTTGGAAATGTACGCCCGCGCTCAAGACATATACGTGAAGG	1031
Oy	5103	TTTTTGGCGGTGTGATCCCTTTATGTGTGGGAGGCGCAATGCTTCTGTACAGTGA	5162

Db	10318	TCCTTGGAGGGGGGTACCCCTTCATGTGGGGAGAGACAAGTTTTCGACAGTGA	103177
Qy	5163	ACACAACAAGTAGAGGCGGTACGTGAGTTCCGCTCAGACTGCATATAGATCAAGCAG	5222
Db	10378	ACACCGAGATGATGAGGGGTACGTGGAATTGTACAGATTTGCGGACTGACACGCGC	104377
Qy	5223	TGCGACTAAAAGTTTCAACAAGCTGTCTGMAAGTCGGCTCGGTATAGTATACGGCACA	5282
Db	10438	AGCGGATTTAAGGTGTCACTACGCGCGCATGAAGATAGAGACTACGTATAGTGTACGGGACA	104977
Qy	5283	CCACCGCGCACCTTGATAGCTTTGTCAATGGCGTCAACGCAAGTTTCTCACGGGACCTTA	5342
Db	10498	CTACACAGTTTCTTAGATGTGTACGTGAACGGAGTCAACACAGGAAGCTTTAAAGACTGA	105577
Qy	5343	AGGTCAATGACAGGGCCGATATCAGCGCTTTTCAACCTTTTGCATTAAGTGTGTCATCA	5402
Db	10558	AAAGTCAATAGCTGACCAATTTACAGATCGTTTACCAATTCGATCAAGAGTGTATATCC	106177
Qy	5403	GAAGGGGCTTGTTTTCAACTACGACTTCCCTGATGATGAGACTATGAACACAGAGCGT	5462
Db	10618	ATCCGCGCGCTGTGTATCAACTATAGACTTCCCGGAATACGAGAGGATGAACACAGAGCGGT	106777
Qy	5463	TGCGCGATTTTCAAGCATCTCTGCGCTGATGCTACACATAGTAGGCCCGCATGACATAC	5522
Db	10678	TTGAGACATTTCAAGCTACTCTCTTGAATAGCAAAATCTCATCGGCACACAGACATTA	107377
Qy	5523	GGCTGTGAAGACCTTCTGTCAAGAACATCAACGCTCCCTACACCCAGAGAGTACAGGT	5582
Db	10738	GACTACTCAAGCCTTCCGCGAAGAGGTGATGTCCCGTACACAGCAGCGCGATCTGAT	107977
Qy	5583	ATGAAATGTGGAAGAACAACTCAGACGACCCCTGCAAGAAACAGCACCATTTTGAATGTA	5642
Db	10798	TCGAGATGTGAAAAAACAATCAAGCCGCGCCACTGCAGAGAAACCGCCCTTCGGGTGCA	108577
Qy	5643	AAATTTGAATGTGAGACCTCTGTGCGAGCGCTTACATCTGTCTTACGGGCAATCTCTATCGA	5702
Db	10858	AGATTTGACATCAATCCGCTTTCGAGCGGTGACGTCTATACGGGAACATTTCCATCTCTTA	109177
Qy	5703	TTGACATCCCTGATATGCAAGCTTTTGTGAGATCATCAGATCAACCAACAAATTTTAGAAGTTA	5762
Db	10918	TCGACATCCCGAAGCGTCTTATACAGACATCAGATGCAACACCTGCTCTAACAGTTCA	109777
Qy	5763	GCTGCACAGTAGCAGACTGCTGATTTATTTCTGCAGACTTTGTGGTCTCTTACATTTACAGT	5822
Db	10978	AATGTGATGTGACAGATGTGCACTTATCAAGCGGACTTTCGGCGGGAATGGCTACCTGTGCAAT	110377
Qy	5823	ACAAAGCTGACAGGAGGGACATTTGTCTCAGTTTCACTTCCACTCAGACAGCTGTTTGA	5882
Db	11038	ATGTATCCGACCGCGAAGGACAAATGCCCTGTATCATTTGCAATTTGCAGCACAGCAACCTCTCC	110977
Qy	5883	AGGAAGGACACACATGTGACTGTGCGGTAGGACGACATTAACATCACTTTTACACATGGA	5942
Db	11098	AAGATGTGACATTCATGTCTCTGGAAGAAAGGACGGGTGACAGTACCTTCACACACCGGCA	111577
Qy	5943	GCCCAACAGCAAAATTTTATAGTTTTCGCTATGCGGCAAGAGTCCAATTGCAATGCTGAT	6002
Db	11158	GCCCAAGGGGAACTTTATGTATTCCTGTGTGTGAAGAAACAATGTCATGCAATGCAAGAT	112177
Qy	6003	GTAAACCAACCGCGCGACCATATATTTGAGAAACCATATTAAGTCGACCAAGATTTCCAG	6062
Db	11218	GCAAAACCAACGCTGACCAATGTGAGAACCCCGCAAAAAATGACCAAGATTTCCAAG	112777
Qy	6063	CGGCAAGTTTCAAAACATCTTGAACGTGGCTGCTTGCATGTTTGGGGAGGATCATCCC	6122
Db	11278	CCGCCATCTCAAAAACCTTCATGAGATTGGCTGTGTTGCCCTTTTTCGGCGGCGCTGTGCGC	113377
Qy	6123	TCATTTGTTAGACATTATAGTGTGTGTGACGCTTATGCTTATTAACACACGTAGAT	6182
Db	11338	TATTAATTATAGACTTATGATTTTTCGTTGCGACATGATGTGCAACAGAAAGAT	113977
Qy	6183	GACTGAGCGGAGACCTGA-CATAGCGGTAAAACTGATGTACTTTCGAGGAAGCGTGT	6241
Db	11398	GACCGCTACGCCCAATGACCCGACACAGCAAAACTGATGTACTTTCGAGGAACGTAGAT	114577

Db 7003 CCACCATCTACTAGGGTACTCGTTTAAATTGGGGCGATGATGAATCCGAATGT 7062
Qy 1900 TCTTAACCGCTGTTTCAACACACTAGTCAATATCATGATGCTGACAGACTACGTTG 1959
Db 7063 TCCTCAACATTTTGTCAACACAGTTTGAATGTGTTTATCGCCAGACAGATCTACGAAG 7122
Qy 1960 AACGGTTAACACAGCTCAGCGTGGCGGCTCTATCGCGCAGATAACATAGTCATGTTG 2019
Db 7123 AGCGGCTTAAACGTCAGATGTCAGCGTTTCATTGGCGACGAAACATCATCATGAGG 7182
Qy 2020 TCGTCTCCGACACTTGTATGGGAGAGATGGCCACTTGGCTGAACATGAGTAAAAA 2079
Db 7183 TAGTATCTGACAAAGAAATGCTGAGAGGTGGCCACCTGGCTCAACATGGAGTTAAGA 7242
Qy 2080 TTATGATGACGATTTTGGTATCAAGACCCCTACTGCTGGGGGATTTATCTGGTGG 2139
Db 7243 TCATGACGACGATCATCGGTGAGAGCCACCTTACTTGGCGCGGATTTATCTTGGCAG 7302
Qy 2140 ACCAGATTAACAGGCAAGCCTGACAGATGCGACACCTCTAAAGAGCTTTTAAAGCTTG 2199
Db 7303 ATTGGTTACTTCCAGAGGTGCCGGGTGGCGGATCCCTGAAAGGCTTTAAAGTTGG 7362
Qy 2200 GAAACCATTTGCGCACTCGATGATACCCAAAGCTGCGACCGCGCGGCACTGCATGATG 2259
Db 7363 GTAAACCCCTCCACCGCAGCAGCAGAAAGACGAAGACGAGCGCTGTGATGATG 7422
Qy 2260 AAGCAATGCAATGAGAAAGAAATTTGAATTAACGAGAGATTAAGAGCGCGTGAATCCA 2319
Db 7423 AAACAAAGGCGGTGTTTAGAGTAAAGTAAACAGGCACTTTAGCACTGGCGGTGACGACCC 7482
Qy 2320 GATACGAGATCATCTGCGACGCGCTGATCATCAGCTCTCTGTCACAGTTAGCCGAAGCG 2379
Db 7483 GGTATGAGTGAACAATTAATTAACCTGTCTACTGGCATTGGAACCTTTGGCCAGAGCA 7542
Qy 2380 TTAAAGACTTCAAGACATTAAGAGGAGCCCAATCACTCTTACGCGCTGACCTTAATAG 2439
Db 7543 AAAGGCAATTCACACCATCAAGAGGGAATTAACATCTTACGCTGCTCTTAATAGT 7602
Qy 2440 TGACGTATTAAGCAGCAGCACTTC-----CCACCGGCGAGATGTTCCATACC 2486
Db 7603 CAGCATAGTACATTTTCATCTGACTTAATCTACAACACACACATGATAGAGATTCCT 7662
Qy 2487 CTCAGCTGAATCTTCCACAGATTTCACCTACAAATCCGATGGCTTACCGAGATCCAAAC 2546
Db 7663 TTAATATGCTGGCGCGCGCCCTTCCCGGCCCACTGCCATGTGAGAGCCGCGAGAA 7722
Qy 2547 CTCTAGCGCGCGCTGAGAGCGGTTTCGCCCCCGCTGGCTGCTCAAAATCGAAGATCTTA 2606
Db 7723 GAGAGCAGGCGGCGCCGATGCTGCCCCGACGAGGCTGGCTTCTCAAAATCCAGCAACTGA 7782
Qy 2607 GAGAGTCGATGATCACTTGAACCTTTCAACACGATCACCTAATCCGCGCGCAAGTCCA- 2665
Db 7783 CCACAGCGGTGAGTCCCTAGTCACTTGAACAGGCAACTGAACCTCAACCCCACTCTCAC 7842
Qy 2666 -----CCGCAAGAAAGAAAGAGTGTCTTAAGCAGAAACCTTACTGAGCTTAAAGA 2720
Db 7843 GCCCGCCACCGCGCCAGAAAGAGCGCGCCCAAGCAACACCGAAGCAAGAAACCA 7902
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGCAAA--GCACGCTTAACCAAGGAAACGACAACTGA 2777
Db 7903 AAACGACAGAGAAAGAAAGAAAGCAACCTGCAAAACCAAAACCGGAAAGAGACGCCA 7962
Qy 2778 TGTGATGAAGTTGGAGTCCGACAAAGACATTTCCGATCATG--CTGAACGCGCAAGTGA 2834
Db 7963 TGGCATTAAGTTGAGGCGGACAGATTTGTCAGCTCAAGAACGAGGACGAGAGTGTCA 8022
Qy 2835 ATGATATGCTGCTGTGTCGAGAGAGGCTGATGAACCACTCCACGTTGAAGAAATA 2894
Db 8023 TCGGGCAGCAGCTGCGCATGAGAGAGATTAATGAACCTTGCACCTGAAAGAGAACCA 8082
Qy 2895 TTGATTAATGACCAATTAGCGCGCGGTGAATTGAAGAGGCTAGCATGATGAGCTTGGAGT 2954

Db 8083 TCGACACCCGTGTGCTATCAAAAGCTCAAAATTTACCAAGTGTGACATATGACATGAGT 8142
Qy 2955 ACGGACAGCTTCCCGAGAAACATGAACACAGCCTGCAGATACACAGCGCAAAACAC 3014
Db 8143 TCGCAGATTTGCGAGTCAACATGAGAGTGAAGCTTCACTTACACAGATGAACCCCG 8202
Qy 3015 CGGGCTTTTACAACTGGACCAAGCGGCGAGTCCAGTATGAAGATGGAATTTTACCGTAC 3074
Db 8203 AAGATTTCTATTAATGCGACCAAGCGGCGTGCAGATATAGTGAAGTATGATTTACATCC 8262
Qy 3075 CGAAGAGTGGGCGGGAAGAGCCGACACCGGAACCGATCTTGGACAAACAGAGCAGAG 3134
Db 8263 CTGGGAGTGAAGAGCAGAGAGACAGCGGTGCTGCATATGATTAATCTCCGCTCGGG 8322
Qy 3135 TTGTGCTATGTTTATGAGGTTGAAGGTCAGCGCGATTAAGCGGCTTTCAGTGCTCA 3194
Db 8323 TTGTGCGATATGCTCTCGGTGCGTGTGATGAAGAAACGAACTGCCCTTTCGCTGCTCA 8382
Qy 3195 CTGGAACCAAGAAAGGGGTGACCATTAAGGATACCCCGAAGTTCGAAACCGTGT-- 3251
Db 8383 CCGTGAATAGTAAAGGAGAAACATTAAGACGACCCCGGAAGGACAGAGAGTGTCCG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGCGTCTTTGGAATGTCACTTCCCATCGCAAAAC 3305
Db 8443 CAGCACACTGCTGACGCGCAATGTGTTGCTCGAAATGTGAGCTTCCCATCGACCGCC 8502
Qy 3306 CACCGCTGTGCTATCACTGACCGCAGAACGAACTGCAAGTGTGCAAGGAACGTGCG 3365
Db 8503 CGCCCAATGCTATACCGCGCAACCTTCAGAGCCCTGACATCTTGAAGGAACGTGA 8562
Qy 3366 ACAATCCAAATTAACAGACAGCTGCTGAGAAAGTCTTGAATGTCATCAC--GCCGCG 3422
Db 8563 ACATGAGGCGCTAGATACCTGCTCATGTCATATTCGGGTGCGGATCTGTGCGAGAA 8622
Qy 3423 CCAAAAGAAATTAACGATATCCGATATCCCTGCTCATGTCATATTCGGGTGCGGATCTGT 3482
Db 8623 GCAAAAGAAAGGCTATGACACATTTTACCTGACAGCCCTTACTTGGGCAATCTGCTGT 8682
Qy 3483 ATGCGAGACCTCAACCGCGGTGTTACGCCCAATTAATTAAGAAACGTGAGGAGCGAAT 3542
Db 8683 ACTGCACATATGCTACCGGTCTTACGCCCTGTTAATGAGAGGCTGTGGAGCAAG 8742
Qy 3543 CTGATGATGATTCGATTAAGATTCAGAGTCTGCGCAAAATTCGGCTTACATCAGCAGCA 3602
Db 8743 CGAGAGATTAACCATAGCATAGACTTCCGCCAATTTGGAATGACACCAACCAAGCGGAG 8802
Qy 3603 CTGCGGATGTCAACAAATTCGTTTACATGCTTTTTCGACACGACCATGATCAAGAGAG 3662
Db 8803 CAGCAAGCGCAAAACAAATACGCTTACATGCTTAAACAGATGACACCGTTAAAGAG 8862
Qy 3663 ACAATGAGAGAAATAGCTATACGACATCTGGAACCTGCGCTGTCTTGGCCCAAG 3722
Db 8863 GCACCTGATGACATCAAGATTTAGCACCTGAGGACCTGTGAGAGGCTTGTGCTCAAG 8922
Qy 3723 GGTACTTCTCTGTAGCTCAATGTCTCTCAGGTTGACAGTGAACCGTCACTGATTCAGAGCG 3782
Db 8923 GATACTTCTCTCGCAAAATTCCTCTCAGGAGACAGGCTTAACGTTAGCATAGTGAAT 8982
Qy 3783 GAGCATCTGAATTCATGACACGCTGAGAAAGAAAGATCAGAGAGAAATTTGCTGTAAG 3842
Db 8983 GCAACTCAGCAACGTCATGTACATGCGCCCGCAAGATTAACAAATTCGTTGAGACGGG 9042
Qy 3843 AGAGTACTTGTTCACACCGGTCCATGGAAGCTGTGAAGTGCACGTTTACATCACT 3902
Db 9043 AAAATATGATCTACCTCCGTTACCGGTAAAAATTCCTTTCGACATGTAGACGCTGC 9102
Qy 3903 TGAAGAGACGCTGCGCGGTACATTAACCATGACAGGCGCAGGCGCAACGCGTTAAGT 3962
Db 9103 TGAAGAAACAACTGCAAGCTCACTATGACACAGGCGCAGACCGCATTTATCAT 9162
Qy 3963 CTTATCTGAGAGAAAGCTGAGGCAAGTGTACATTAACCACTTTTGGCAAGAACGTCA 4022
Db 9163 CTTACTGGAAGAAATCATCAGGAAAGTTTACGCAAGCCGCACTCTGGAAGAACATTTA 9222

OY	4023	CCATCAAGATGAATGATGGGACCTACAGCA	CAGTATTCGTGAGACAGCGAAGCAAGATGA	4082
Db	9223	CGATATGATGGAAATGGCGGACCTACAAAG	CCGGAACCGTTTCGACCCGACCCGAATGA	9282
OY	4083	ACGGCTCAGCTAAAGCAAAACAGTCACTTC	CTTACAAAGACGCAAAAGAAATGGGTCT	4142
Db	9283	CTGGTTACACCCCATCAGACAGTGCGTGC	CTTTAAGAGCGACCAAGAGTGGTCT	9342
OY	4143	TTCAACTGGCGGATCTTATTAGGACACAGA	CCACTTCAGTGCAGGTAAATTGCACATTC	4202
Db	9343	TCAACTACCGGACCTTATCAGCATGACAC	ACACACACACGCGCCAGGAAATTCGATTTC	9402
OY	4203	CATTCCGCTTGACACCGACAGTCTGCGCC	GGTTCGGTTAGTCCACAGCCCTACAGTCA	4262
Db	9403	CTTTCAAGTTGATCCGAGTACTGTGATG	ATGTTCTCTTGCCACGCGCGGAATGTATAC	9462
OY	4263	AGTGTTCAAAAGGACATCACCTCCACTG	ACTGCAATGCGACCAATTTGCTGACACGA	4322
Db	9463	ATGGCTTTAAACATCAGCTCCCAATTAT	ATACAGACCACTTGACATTGTCTACACCA	9522
OY	4323	GAAATTTGGGCTGCGACAGACGCAACGA	ATGATATGATTTACAGGTTCTACATCCAGA	4382
Db	9523	GGAACATAGGGGCAAAACCGGAACCAAC	CACTAATGTGATCTGTGAAAGACGTTAGAA	9582
OY	4383	ATTTTTCGTGGGCGAGAGGCGTGGAGTA	TCATATGGGTTACATGAAACAGTCAAG	4442
Db	9583	ACTTCACCGTCGACCGAGATGCGCTG	GAATATACATATGGGAAATCATGAGCCAGT	9642
OY	4443	TCGTGGCCCAAGAGTCGGCACACAGGCG	ACCACTGATATGATGCGCGCATGAGATCATCC	4502
Db	9643	TCATATGCCAAAGATCGACACAGGAAAC	CCCTCAGATGGCCACAGAAATGTACAC	9702
OY	4503	ACATATTCATCCGACATCCAGTCTAC	CTGTCACTTATGTCGTGTGGTGCCTTGCA	4562
Db	9703	ATTATCTACATCCGCATCTCTGTGACCA	TCATTTAGCCGTGATACGCTACCGTGGCA	9762
OY	4563	TCCTGTAGGCACTGATCATCAGACCTTG	ATGCGCAAAAGCAAGAAAGACATGCTGTA	4622
Db	9763	TGATGATTTGGGTTAACTGTGACAGTGT	ATATGTGCTGTAAAGCGCCGCTGAGTGGTA	9822
OY	4623	CGCATACGCGCTTGACACCGAACGCA	ACGTTACCCACAGATTAGGGTTTGTGCTGA	4682
Db	9823	CGCATACGCGCTTGACACCGAACGCG	TAATCCCACTTGCTGCGCACTCTTGTGCTGG	9882
OY	4683	TTCCGCGCAACCAACGCTGAAACATTTG	AGAAACCTTTGAAACCACTCTGTGTTTAAACAC	4742
Db	9883	TTTAGTGGGCCAATGCTGAAACGTTAC	CGAGACATGATTACTTGTGGTGAACAGTC	9942
OY	4743	AACCGTTTCTCTGGGCAAGTTGTGATTC	CTCTGCGACGCTTGTATTCTGTTCGGT	4802
Db	9943	AGCGTTCTCTGGGTCCAGTTGTGCAT	CACTTTGGCGCTTTCATCGTTCAATTCGGCT	10002
OY	4803	GCTTTATGCTGATACCTTTTATTTG	GTACAGGCGCTGCGTGGGGAAAGTATGAGC	4862
Db	10003	GCTCTCTCTGCTGCTGCTCTTTTAT	GTGTGTGCGGCGCTTACCTGTGGAAAGTATGAGC	10062
OY	4863	CCTTGCAACATGCGACCACTGTGCAAA	TGTTCCGGGATCCCGATTAAGCGCTTGCTGC	4922
Db	10063	CCTACGAACATGCGACCACTGTTCCAA	ATATGTGCACAGATACGTTAAGGCACTTGTTG	10122
OY	4923	AACGCGCAGTTACGCGCACTTAACTG	AGATCAACGTCCTCATGTGGAATTAAACG	4982
Db	10123	AAAGGGCAGGTTAGCCCGCTCAATTTG	GAAGATCACTGTCACTGCTCGGAGTTTTCG	10182
OY	4983	CTTCAACTAACAGAGATAGTGACCTG	CAAAATTTCCACACAGTCATCTCTCAACCAAG	5042
Db	10183	CTTCAACCAACAGAGATATCATTTAC	CTGCAAAATTCACACATGTGTGCCCTCCCCA	10242
OY	5043	TTAAATGCTGCGGTCCTCGATG	CAAGCATCTCTAAAGCGGATTTACATGTCGCGG	5102
Db	10243	TCAATGCTGCGCTCTCTTGGAATGTAC	GCGCGCGCTCATGTGACATATACCTTGCAAG	10302

QY	5103	TTTTTGGCGGTGTGATCCCTTTCACTGTGGGAGGCGCAATGCTTCTGTGACAGTGA	5162
Db	10303	TCTTTCGGAGGGGTCTACCCCTTTATGTGGGAGGAGCGCAATGTTTTTGGCAAGTGA	1036
QY	5163	ACAACAACCTAGTGAAGGCGTATCGAGTTCGCTCAGACTGTCACTATAGATACCGCAG	5222
Db	10363	ACAGCGAATATAGTGAAGGCGTATCGAATTTGTACAGAGATTGGCGCTCTGACACCGCC	1042
QY	5223	TGCGACTAAAGTTCAACAGCTGCTCTGAAGTGGCGCTGCGTATATGATATACGGCAACA	5282
Db	10423	AGGCGATTTAAGTGCACACTGCGCGCATATGAAGAAAGTAGAGACTGCGTATTTGTGACGGGAACA	1048
QY	5283	CCACCGCGCACTCTGATATCGTTGTCAATGGCGCTCAGCCAGAGTTTCTCAGCGGACCTGA	5342
Db	10483	CTACCGAGTTTCTTGAATGTGTACGTGAACGGAGTCAACCAAGGAACGCTCTTAAACCTTGA	1054
QY	5343	AGGCTATGCAAGGCGCATACACCGGCTTTTACCGCTTGAACCTATTAAGTGGTCAACA	5402
Db	10543	AAGCTATAGCTGCAACCAATTTCAACATCGTTTACGCAATTCGATCATTAAGGTGCTTATCC	1060
QY	5403	GAAAGGGGCTTGTTCACACTACCACTTCCCTGAGTATGAGACTATGAACACAGAGCGT	5462
Db	10603	ATCCGCGCTGTGTACAACTATACACTTCCCGAATATGAGGCGATTAACACAGAGCGT	1066
QY	5463	TCGGCGATTTCAAGCATCTCGTCTGATGTCAAGACATATGATGCCCGCATGACTATAC	5522
Db	10663	TTGGAGACATTTCAAGCTACCTCTTGTGACTAGCAAGGATCTCATGCGCAGCAAGACATTA	1072
QY	5523	GGCTGCTGAACCTCTTGTGCAAGAACATCCACGTCCTTACACCAAGCAAGTATCAGAGGT	5582
Db	10723	GGCTACTCAACCTTCGCGCAAGACGTGCATGTCCTGTATACCAAGGCTCTCATCAGAT	1078
QY	5583	ATGAATGTGNAAGAACAACTCAGAGGACCCCTGCAAGAAACAGCAACAATTTGGATGTA	5642
Db	10783	TTGAGATGTGAAABAAACAACACAGCGCGCCACATCGAGAAACCGCACTTTGGGGTGTGA	1084
QY	5643	AAATTTGAAGTGAAGCTCTGCGAGCGCTTAACTGTGCTTACGGGCAATCTCTTATCTGCA	5702
Db	10843	AGATTTGACGTAAATCCGCTCCGACCGGTGAGCTGTATATCGGGAAACATTTCCCTTCTA	1090
QY	5703	TTTGACATCCCGATGACAGCTTTTGTGATGATCATAGAAATCAACAAATTTTAAGTTGA	5762
Db	10903	TTGACATCCCGAAGCTGCTCTTATACAGAAATCAGATGACACCACTGGTCTCAACAGTCA	1096
QY	5763	GCTGCACTGATGACAGACTCATTTATTTCTGAGACTTTGGTGTCTCTTAACATTACATGAT	5822
Db	10963	AATGGAAGTCAGTGAGTGCATTTATTCAGAGACTTCGGGGGATGGCCACCTGCGAGT	1102
QY	5823	ACAAAGCTGACAGGAGGAGCAATTTGCTCAAGTTCACTCCACTCCACGACAGCTGTTTGA	5882
Db	11023	ATGTATCCGACCGGGAAGGTCAATGCCCGGTACATTTGCGACTTCGAGCAAGCAACTCTCC	1108
QY	5883	AGGAAGGCAACCACTATGTGACTGCCGTAGGCAAGCATTAACATTCATTTTATAGCAATCGA	5942
Db	11083	AAGAGTCAACAGTCACTGTCTCTGGAAGAAAGGAGCGGAGCAAGTACATTTTAGCACCCGCA	1114
QY	5943	GCCCAAGCAAGAAATTTATAGTTTTCGCTATCGCGCAAGAAAGTCAACTGCAATGCTGAAT	6002
Db	11143	GTCACAGGCGAATCTTATCTATCTGCTGTGTGGAAAGAAAGAACAACTATGCAATGCAAAAT	1120
QY	6003	GTAACCAACCGGCGGACCACTAAATTTGAGAGAACCAATAAAGTGCACAAAGATTCGAG	6062
Db	11203	GTAACCAACCAAGCTGACCATATCTGAGACACCCGCACAAAAATGACCAAGAAATTTCAAG	1126
QY	6063	CGGCAAGTTCCAAAACATCTTGGAACTGGCTGCTTGCATGTTTGGGGAGACATCATCC	6122
Db	11263	CGGCAATCTCAAAAACATCATGAGGATGGCGCTGTTGCCCTTTTGGGGGCGCTCTGTCGC	1132
QY	6123	TCATTTGTTGTAAGCTTATAGTGTGGGTCTGCTGACCTGTATGACTTAATAACACAGCTAGAT	6182
Db	11323	TATTTATTTATGACCTTATGATTTTTTGTCTTGACAGATATGCTGATCTAGCACAGAAAGAT	1138
QY	6183	GACTGAGCGCGGACACTGA-CATAGCGGTAATACTGATATCTTCGAGAGACGCTGT	6241

Db	11383	GACCGCTACGCCCCCAATGATCCGACGACAAACTCGATGTACTTCGAGGAACTGATGT	11442
Oy	6242	GCATATGCCACGCGCGCT	6261
Db	11443	GCATATGCAATCAGGCTGGT	11462

Search completed: November 16, 2003, 09:06:06
Job time : 380 secs

